UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
AD-A208 590 DTIC ELECTE JUN 06 1989 D	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited. (2)
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 89-0729
6a. NAME OF PERFORMING ORGANIZATION: Columbia University	6b. OFFICE SYMBOL (If applicable):
6c. ADDRESS (City, State, and ZIP Code): Department of Mathematics, Box 20, Lou Monorial Lab. New York, NY 10027	7a. NAME OF MONITORING ORGANIZATION: AFOSR
	7b. ADDRESS (City, State, and ZIP Code): Building 410, Bolling, AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NM
	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-87-0117
8c. ADDRESS (City, State, and ZIP Code): Building 410, Bolling, AFB DC 20332-6448	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	A4	

11. TITLE (Include Security Classification): Differential Equations, Related Problems of Pade Approximations and Computer Applications

12. PERSONAL AUTHOR(S): D.V. Chudnovsky and G.V. Chudnovsky

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 1 JAN 87 TO 31 DEC 88		5

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

During the past period of the Grant, ~~our~~ this work focused on the study of analytic, arithmetic and algorithmic properties of differential equations applied to solutions of problems in theoretical mathematics, mathematical and theoretical physics, numerical methods and computer science. Keywords:

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. Arje Nachman	22b. TELEPHONE (Include Area Code): (202) 767-4939 22c. OFFICE SYMBOL: NM

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED

89 6 06 074

LIST OF PUBLICATIONS OF D.V. CHUDNOVSKY AND G.V. CHUDNOVSKY FOR 1986/88 SUPPORTED BY U.S. AIR FORCE

1. A random walk in higher arithmetic. Advances in Applied Mathematics, v. 7 (1986), 101-122.

2. Sequences of numbers generated by addition in formal groups and new primality and factorization tests. Advances in Applied Mathematics, v. 7 (1986), 187-237.

3. On expansion of algebraic functions in power and Puiseux series, Parts I and II. J. of Complexitites, v. 2 (1986), 271-294, and v.3 (1987), 1-25.

4. Computer assisted number theory. IBM Research Report, RC 12030, 7/23/86, 67pp. Lecture Notes in Mathematics, Springer, N.Y., v. 1240, 1987, 1-68.

5. Elliptic modular functions and elliptic genera. Topology, v.27 (1988), 163-170.

6. Algebraic compliexities and algebraic curves over finite fields. Proc. Natl. Acad. Sci. U.S.A., v.84 (1987), 1739-1743.

7. Algebraic complexities and algebraic curves over finite fields. IBM Research Report, RC 12605, 3/24/87, 50pp; J. of Complexity, v. 4, No.3 (1988), 285-316.

8. Elliptic formal groups over $\mathbf{Z}$ and $\mathbf{F}_p$ in applications to nnmber theory, computer science and topology. Lecture Notes in Mathematics, Springer, v. 1326 (1988), 11-54.

9. Computer Algebra in the Service of Mathematical Physics and Number Theory. Proceedings of 1986 Standord Conference "Computers and Mathematics", (to appear).

10. Transcendental methods and theta-functions, Proc. Symp. Pure Math., Proc. AMS Summer Schools in Maine, 1988 (to appear).

11. Approximations and complex multiplication according to Ramanujan, in Proceedings of the Ramanujan Centenary Conference, Academic Press, 1988, 375-472.

12. Regular graphs with small diameter as models for interconnection networks, Proc. Third Intl. Conference on Supercomputing, Boston, 1988, v. III, 232-239 (with M. Denneau).

13. A design of general purpose number-theoretic computer, ibid, v. II, 498-492 (with M. Denneau, S. Younis).

Final Technical Report

Differential Equations, Related Problems of Pade Approximations and Computer Applications

D.V. Chudnovsky, G.V. Chudnovsky

Department of Mathematics

Columbia University

1988

1. Study of Arithmetic Properties of Linear Differential Equations

For the period of the Grant we conducted examinations of arithmetic and geometric properties of linear differential equations, using methods of Padé approximations. These studies were aimed at detection of special arithmetic properties of solutions of these equations that can be used to study diophantine approximations of classical constants. Recently, we concentrated our work on examination of arithmetic and geometric properties of linear differential equations that distinguish the deformation equations (Picard-Fuchs equations) from other linear differential equations.

About 20 years ago it was discovered and established that linear differential equations having a geometric sense, like the Picard-Fuchs equations satisfied by the variation of periods, possess strong arithmetic properties (global nilpotence, action of the Frobenius, Fuchsianity,... etc). We refer to the review [1]. Since then it was suspected, that in a certain sense, the converse is true too. We found new theoretic and experimental evidence that show that linear differential equations with arithmetic (or integrality) properties of their solutions arise from geometry, or, precisely, correspond to deformations of period structure of algebraic manifolds. A variety of problems from diophantine geometry arise here, including the irrationality and diophantine approximations to constants of classical analysis.

We start with the definition of the basic class of function. In his seminal paper [3] on diophantine approximations, Siegel defined and targeted for future studies two classes of functions satisfying linear differential equations and given by power series expansions in x, for the values of which one can establish general theorems on irrationality, transcendence and on the measure of linear independence. These two classes are classes of E-functions and G-functions, that command attention of modern diophantine approximations research. The study of E-functions, started by Siegel [3-4], were significantly advanced since then by many researchers, see particularly [5-6]. We would like to mention in this connection that only relatively recently we have proved results on the best possible measure of diophantine approximations of values of E- functions at rational points [7]. These results present an ultimate effective version of the Schmidt theorem [9] for the values of E-functions. In algebraic geometry and analysis, however, most of the interesting functions are analytic only in the finite part of the complex plane and have much better p- adic convergence properties. Among these the G-functions play the crucial arithmetic role.

*Definition*1.1. A function $f(x)$ with the expansion at $x = 0$:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

is called a G-function, if $f(x)$ satisfies a linear differential equation over $\overline{\mathbf{Q}}(x)$, if coefficients a_n are algebraic numbers, and if there is a constant $C > 1$ such that for all $n \geq 0$ the sizes

of coefficients a_n (i.e. the maximum of absolute values of a_n and all its conjugates) and the common denominators $\mathrm{den}\{a_0, \ldots, a_n\}$ are bounded by C^n.

Siegel introduced this class of G-functions and put forward a program to prove the linear independence theorems for values of G-functions at algebraic points near the origin. Unfortunately, Seigel [3] never explicitly proved general theorems for G-functions, instead presenting examples of such theorems and presenting an outline of a theory that could be constructed similar to the theory of E-functions. Progress in the later study of G-functions becomes heavily dependant on additional very restrictive global "geometric" conditions, formulated for the first time by Galočhkin [9], that demand that the G-function property be shared by an expansion of any other solution of a differential equation satisfied by $f(x)$ at any algebraic point. We called these global conditions in [10], [11] the (G, C)-conditions. Previously known results on G- functions rely exclusively on these (G, C)-conditions, or on equivalent ones [9], [12], [13], [11]. These global conditions can also be reformulated [12] in terms of the p-adic "overconvergence at a generic point" of solutions of a linear differential equation satisfied by $f(x)$.

In [14] and [15] we had proved the general linear independence results for values of arbitrary G-functions at algebraic points (close to the origin), without any additional conditions. These results materialize Siegel's program after some 55 years. Also in [14], [15] proofs of results on the absence of algebraic relations are presented. It is more important, however, that we have proved the strong (G, C)-property for arbitrary G- functions [20]. This result, connected with our study of the Grothendieck conjecture, implies, e.g. that all previous results on G-function theory, proved under very restrictive conditions, are unconditionally valid for all G-functions. To describe our result, and the (G, C)-function conditions, one needs the definition of the p-curvature.

We consider a system of matrix first order linear differential equations over $\mathbf{Q}(x)$, satisfied by functions $f_i(x) : i = 1, \ldots, n$:

$$df_i(x)/dx = \sum_{j=1}^{n} A_{i,j}(x) f_j(x), \tag{1.1}$$

for $A_{i,j}(x) \in \mathbf{Q}(x) : i, j = 1, \ldots n$. Rewriting the system (1.1) in the matrix form

$$df^f/dx = Af^f; A \in M_n(\mathbf{Q}(x)),$$

one can introduce the p-curvature operators Ψ_p, associated with the system (1.1). The p-curvature operators Ψ_p are defined for a prime p, as

$$\Psi_p = (d/dx - A)^p (\mathrm{mod} p).$$

Then Ψ_p is a linear operator that can be represented as $\Psi_p = -A_p(\mathrm{mod} p)$, where one defines for $m \geq 0$,

$$(d/dx)^m \equiv A_m(\mathrm{mod}\mathbf{Q}(x)[d/dx](d/dx - A)). \tag{1.2}$$

Let $D(x)$ be a polynomial from $\mathbf{Z}[x]$ that is the denominator of A, i.e. $D(x)A_{i,j}(x)$ is a polynomial in $\mathbf{Z}[x]$ for $i, j+1, \ldots, n$. The *(G,C)-function condition* [10]- [11] of (1.1) means

that (1.1) is satisfied by a system $(f_1(x),\dots,f_n(x))$ of G-functions, and that there exists a constant $C_2 > 1$, such that for any N, the common denominator of all coefficients of all polynomial entries of matrices $D(x)^m A_m(x)/m! : m = 0,\dots,N$, is growing not faster than C_2^N. With this conditions is closely related a *global nilpotence* condition [15-18] stating that the matrices Ψ_p are nilpotent for almost all primes p. The (G,C)-condition implies the global nilpotence condition.

In [15] we proved the global nilpotence (and the (G,C)- function condition) of linear differential equations having a G-function solution. To prove this result we used Padé approximants of the second kind.

*Theorem*1.2. Let $f_1(x),\dots,f_n(x)$ be a system of G-functions, satisfying a system of first order linear differential equations (1.1) over $\overline{\mathbf{Q}}(x)$. If $f_1(x),\dots f_n(x)$ are linearly independent over $\overline{\mathbf{Q}}(x)$, then the system (1.1) satisfies a (G,C)-function condition and is globally nilpotent.Any solution of (1.1) with algebraic coefficients in Taylor expansions is a G-function.

Let $Ly = 0$ be a linear differential equation of order n over $\overline{\mathbf{Q}}(x)$ satisfied by a G-function $y(x)$, and $y(x)$ does not satisfy a linear differential equation over $\overline{\mathbf{Q}}(x)$ of order $< n$. Then the equation $Ly = 0$ is globally nilpotent, and all solutions of the equation $Ly = 0$ with algebraic initial conditions at an algebraic point $x = x_0$ have G-function expansions at $x = x_0$.

Our main tool in the study of the Grothendieck conjecture, and in the current study of globally nilpotent equations is the analytic method of Padé, and more general algebraic approximations to functions satisfying nontrivial complex analytic and arithmetic (p-adic) conditions. The corresponding group of results can be considered as a certain "local-global" principle. According to this principle, algebraicity of a function occurs whenever one has a near integrality of coefficients of power series expansion—*local conditions*, coupled with the assumptions of the analytic continuation (controlled growth) of an expanded function in the complex plane (or its Riemann surface)—a global, *archimedian* condition.

To prove the algebraicity of an integral expansion of an analytic function, only assumptions on a uniformization of this function have to be made. Our results from [19] and [20] were proved in the multidimensional case as well, to include the class of functions, uniformized by Jacobi's theta-functions (e.g. integrals of the third kind on an arbitrary Riemann surface). Moreover, our result includes "the nearly-integral" expansions, when the denominators grow slower than a typical factorial $n!$ denominator. E.g., one of the results [19-20], shows that $g+1$ functions in g variables having nearly integral power series expansions at $\overline{x} = \overline{0}$ and uniformized near $\overline{x} = \overline{0}$ by meromorphic functions of finite order of growth *are algebraically dependent*.

Among the applications of this result are some partial but effective results on the Tate conjecture on the bijectivity of a map $\mathrm{Hom}(A,B) \otimes \mathbf{Z}_l \to \mathrm{Hom}(T_l(A),T_l(B))$ for Abelian varieties A and B over algebraic number fields. The Tate conjecture for elliptic curves was left open by Serre [21] in the case when A and B have integral invariants but no complex multiplication. Finally Faltings [22] proved (ineffectively) the finiteness of the isogeny classes for arbitrary Abelian varieties, solving the Tate, Schafarevich and Mordell conjectures. We proved the effective version of the Tate conjecture for elliptic curves using

only [20] and the Honda [23] criterion of isogeny of elliptic curves in terms of the logarithms of their formal groups over $\mathbf{F}_p$:

*Corollary*1.3. If two elliptic curves $E_1/\mathbf{Q}$ and $E_2/\mathbf{Q}$ have the same number of points mod p for almost all p, (or even for almost all p below an effective bound), then $E_1/\mathbf{Q}$ and $E_2/\mathbf{Q}$ are isogeneous over $\mathbf{Q}$.

The demand to have a meromorphic uniformization of is a restrictive one, and, combined with nearly integrality condition, limit us pretty much to class of functions uniformized by θ-functions. However, results similar to [19-20] can be proved for solutions of linear differential equations uniformized by special Fuchsian groups. Namely, it is true for functions with nearly integral coefficients that satisfy linear differential equations, whose monodromy group is up to the conjugation a subgroup of $GL_n(\overline{\mathbf{Q}})$

We have applied the Padé approximations methods used in diophantine approximations and G-functions, to the Grothendieck conjecture that determines the global (monodromy) properties of a linear differential equation in terms of reductions (mod p) of this differential equation.

TheGrothendieckConjecture. If a scalar linear differential equation of order n over $\mathbf{Q}(x)$ has n solutions (mod p) in $\mathbf{F}_p(x)$, linearly independent over $\mathbf{F}_p(x^p)$, for almost all prime numbers p, then this linear differential equation has only algebraic function solutions.

Equivalently, if a matrix system (1.1) of differential equations over $\overline{\mathbf{Q}}(x)$ has a zero p-curvature $\Psi_p = 0$ for almost all p, then this system (1.1) has algebraic function solutions only.

Using our local-global algebracity results, we solved the Grothendieck conjecture for large classes of equations, including equations, solutions of which can be parametrized by means of multidimensional theta-functions. To the class of these equations belong equations of rank one over arbitrary (finite) Riemann surfaces [20]:

*Theorem*1.4. Any rank one linear differential equation over an algebraic curve, i.e. a first-order equation with algebraic function coefficients, satisfies the Grothendieck conjecture. Namely, if Γ is an algebraic curve (given by the equation $Q(z,w) = 0$) over $\overline{\mathbf{Q}}$, and if the rank one equation

$$\frac{dF}{F} = \omega(z,w)dz \tag{1.3}$$

over $\overline{\mathbf{Q}}(\Gamma)$ (for an Abelian differential ωdz on Γ) is globally nilpotent, then all solutions of (1.3) are algebraic functions.

The relationship of the p-curvature operators with the monodromy (Galois) group of a differential equation is extremely interesting. Our methods, involving various generalizations of Padé approximations, allow us to prove the Grothendieck conjecture for a larger class of differential equations, when additional information on a monodromy group is available. For example, if a second order linear differential equation has a commutative monodromy group, then this equation satisfies the Grothendieck conjecture (the Lamé' equation with an *integral* parameter n belongs to this class). A technique from [27] using a random walk method (in the free group, corresponding to the representation of a full

modular group $SL_2(\mathbf{Z})$), allowed us to treat crucially important class of equations. The random walk technique used in [27] is the extension of the earlier work [28]. Among the results of [27] is a

*Theorem*1.5 Let $Ly = 0$ be an n-th order linear differential equation over $\overline{\mathbf{Q}}(x)$ satisfying the assumptions of the Grothendieck conjecture. If the monodromy group of $Ly = 0$ is up to a conjugation a subgroup of $GL_n(\overline{\mathbf{Q}})$, then all solutions of $Ly = 0$ are algebraic functions.

Our results on the Grothendieck conjecture are effective (i.e. one has to examine only a finite set of primes p), and can be used in various algorithms, including algorithms that determine the reduction of Abelian integrals to elementary functions, see [24].

Various applications of our new methods to the proof of transcendence of numbers appearing as elements in the monodromy matrices of linear differential equations are possible. One of our results is the following [28]:

*Theorem*1.6. Let us look at a G-function solution $(f_1(x), \ldots, f_n(x))$ of (1.1) with algebraic initial conditions at a non-singular point of (1.1). Then its analytic continuations along (the basis of) all possible paths leads to at least one transcendental number.

Another application is the Matthews problem [29] on indefinite integration of algebraic functions. Let Γ be a curve over $\mathbf{Q}$ and D be any derivation of a function field $\mathbf{Q}(\Gamma)$. Let $f \in \mathbf{Q}(\Gamma)$ be such that for almost all p, we can find g in $\mathbf{F}_p(\Gamma \mathrm{mod} p)$ such that, mod p, $f = Dg$. Is it true then, that

$$f = Dg$$

for some g in $\mathbf{Q}(\Gamma)$? (I.e. if an integral is locally algebraic at the same field, is it globally algebraic?) The answer is yes.

While the Grothendieck conjecture tries to describe the equations, *all* solutions of which are nearly integral, it is more important to find out which equations possess G-function solutions. These equations, in all known cases, admit the action of Frobenius.

Next, all evidence points towards the conjecture that the globally nilpotent equations are only those equations that are reducible to Picard-Fuchs equations (i.e. equations satisfied by Abelian integrals and their periods depending on a parameter).

Our results on G-functions allow us to represent this conjecture even in a more fascinating form, see [37]:

SiegelConjecture. Let $y(x) = \sum_{N=0}^{\infty} c_N x^N$ be a G-function (i.e. the sizes of c_N and the common denominators of $\{c_0, \ldots, c_N\}$ grow not faster than the geometric progression in N).If $y(x)$ satisfies a linear differential equation over $\overline{\mathbf{Q}}(x)$ of order n (but not of order $n-1$), then the corresponding equation is reducible to Picard-Fuchs equations. In this case $y(x)$ can be expressed in terms of multiple integrals of algebraic functions.

Siegel, in fact, put forward a conjecture which is, in a sense, stronger than the one given above. To formulate Siegel's conjecture we have to look again at his E-functions defined in the same paper [3] of 1929, where G-functions were defined, and where Siegel's theorem on integral points was proved. We remind that a function $f(x) = \sum_{N=0}^{\infty} \frac{c_N}{N!} x^N$

is called an E-function is c_N are algebraic numbers and for any $\epsilon > 0$, the size $|\overline{c_N}|$ of c_N and the common denominator of $\{c_0, \ldots, c_N\}$ is bounded by $O(N!^{\epsilon})$; one also assumes that $f(x)$ satisfies a linear differential equation over $\overline{\mathbf{Q}}(x)$. Siegel showed that the class of E- functions is a ring closed under differentiation and integration. Siegel also studied the hypergeometric functions

$$ {}_mF_n\left(\begin{matrix}a_1,\ldots,a_m\\ b_1,\ldots,b_n\end{matrix}\middle|\lambda x\right) $$

for algebraic $\lambda \neq 0$, *rational* parameters $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ and $m \leq n$. These functions he called hypergeometric E=functions and suggested in [4] all E-functions can be constructed from hypergeometric E-functions.

Looking at the (inverse) Laplace transform of $f(x)$, we see that Siegel's conjecture translates into a conjecture on G-function structure stronger than Siegel's conjecture given above. Indeed, it would seem that all Picard-Fuchs equations might be expressed in terms of generalized hypergeometric functions!

This stronger conjecture is not entirely without merit; e.g. one can reduce linear differential equations over $\overline{\mathbf{Q}}(x)$, satisfied by G-functions to higher order equations over $\overline{\mathbf{Q}}(x)$ with regular singularities at $x = 0, 1, \infty$ only–(like the generalized hypergeometric ones) cf. [23].

We are unable so far to give a positive answer to the conjecture that all arithmetically interesting ($G-$)functions are solutions of Picard-Fuchs equations. Nevertheless, in some cases we can prove that this conjecture is correct. For now our efforts are limited to the second order equations (which provides with an extremely rich class of functions).

It is very likely that, at least for the second order equations, the Siegel conjecture can be proved *modulo* the full Grothendieck conjecture.

2. P-adic properties and P-adic Spectrum of Linear Differential Equations

First interesting applications of arithmetic studies of linear differential equations occur for second order equations with 4 regular singularities,(Heun equations), and, particularly, for the Lamé equation:

$$y'' + \frac{1}{2}\{\frac{1}{x} + \frac{1}{x-1} + \frac{1}{x-a}\}y' + \frac{B - n(n+1)x}{4x(x-1)(x-a)}y = 0 \qquad (2.1)$$

(depending on n and on accessory parameters B).

A more familiar form of the Lamé equation is the transcendental one (with the change of variables: $a = k^{-2}$, $x = (sn(u,k))^2$) [26]:

$$\frac{d^2y}{du^2} + k^2 \cdot \{B - n(n+1)sn^2(u,k)\}y = 0 \qquad (2.2)$$

or

$$\frac{d^2y}{du^2} + \{H - n(n+1)\mathcal{P}(u)\}y = 0. \qquad (2.3)$$

For the Lamé equation (2.3) with an integral n defined over $\overline{\mathbf{Q}}$ (i.e. $a \in \overline{\mathbf{Q}}$ and $B \in \overline{\mathbf{Q}}$), our local-global principle of algebraicity can easily solve the Grothendieck conjecture [22], [26]. We proved in [22]:

Theorem2.1. For integer $n \geq 0$ the Lamé equation has zero p - curvature for almost all p if and only if all its solutions are algebraic functions. The Lamé equation is globally nilpotent for $2n+1$ values of B : $B = B_n^m$ - ends of lacunae of spectrum of (2.3).

For all other values of B, the global nilpotence of the Lamé equation with integral n over $\overline{\mathbf{Q}}$ is equaivalent to the algebraicity of all solutions of (2.3).

The possibility of all algebraic solutions of (2.3) with $B \neq B_n^m$ was shown by Baldassari (Dwork). It can correspond only to special torsion points $H = \mathcal{P}(u_0)$ for special elliptic curves.

For nonintegral n no simple uniformization of solutions of Lamé equation exits. Moreover, Lamé equations themselves provide the key to several interesting uniformization problems. An outstanding Lamé equation is that with $n = -1/2$. This equation determines the uniformization of the punctured tori. If one starts with a tori corresponding to an elliptic curve $y^2 = P_3(x)$, then the function inverse to the automorphic function, uniformizing the tori arises from the ratio of two solutions of the Lamé equation with $n = -1/2$.

$$P_3(x)y'' + \frac{1}{2}P_3'(x)y' + \frac{x+C}{16}y = 0, \qquad (2.4)$$

or

$$\frac{d^2y}{du^2} + [H + \frac{1}{4}\mathcal{P}(u)]y = 0.$$

If $P_3(x) = x(x-1)(x-a)$ (i.e. the singularities are at $x = 0$, a and ∞), then the monodromy group of (2.4) is determined by 3 traces $x = \mathrm{tr}(M_0M_1)$, $y = \mathrm{tr}(M_0M_a)$, $z = \mathrm{tr}(M_1M_a)$. Here M is a monodromy matrix in a fixed basis corresponding to a simple loop around the singularity a. These traces satisfy a single Fricke identity :

$$x^2 + y^2 + z^2 - xyz = 0.$$

The (Poincaré) uniformization case is that, when in (2.4) the monodromy group can be represented by real 2×2 matrices. There exists a single value of the accessory parameter C for which the uniformization takes place.

AlgebraicityProblem[27]. Let an elliptic curve be defined over $\overline{\mathbf{Q}}$ (i.e. $a \in \overline{\mathbf{Q}}$). Is it true that the corresponding (uniformizing) accessory parameter C is algebraic? Is the corresponding Fuchsian group a subgroup of $GL_2(\overline{\mathbf{Q}})$ (i.e. x, y and z are algebraic)? Conversely, if x, y and z are algebraic corresponding to a Fuchsian subgroup of $SL_2(\overline{\mathbf{Q}})$, is it true that the uniformized tori is defined over $\overline{\mathbf{Q}}$?

Extensive multiprecision computations, we first reported in [27], of accessory parameters showed rather bleak prospect for algebraicity in the accessory parameter problems.

Namely, as it emerged, there are only 4 (classes of isomorphisms of) elliptic curves defined over $\overline{\mathbf{Q}}$, for which the values of uniformizing accessory parameters are algebraic. These 4 classes of algebraic curves are displayed below in view of their arithmetic importance.

We want to explain here that the attention to the arithmetic properties of the Lamé equation with $n = -1/2$ arose shortly after Apéry's proof of the irrationality of $\varsigma(2)$ and $\varsigma(3)$. His proof, 1978, was soon translated into assertions of integrality of power series expansions of certain linear differential equations.

To look at these differential equations we will make use of the classical equivalence between the punctured tori problem and that of 4 punctures on the Riemann sphere. For differential equations this means Halphen's algebraic transformation between the Lamé equation (2.4) with $n = -1/2$ for $P(x) = x(x-1)(x-a)$, and the Heun equation with zero-differences of exponents at all singularities:

$$P(x)y'' + P'(x)y' + (x + H)y = 0. \tag{2.5}$$

$$C = 4H + (1 + a).$$

We have already stated that there are 4 Lamé equations with $n = -1/2$ (up to Möbius transformations) for which the value of the accessory parameter is known explicitly and is algebraic. These are 4 cases when the Fricke equation

$$x^2 + y^2 + z^2 = xyz,$$

with $0 \le x \le y \le z \le xy/2$, has solutions, whose squares are integers. It is in these 4 cases, when the corresponding Fuchsian group uniformizing the punctured tori is the congruence (arithmetic) subgroup, see references in [23], [27].

Let us look at these 4 cases, writing down the corresponding equation (2.5):

1) $x(x^2-1)y''+(3x^2-1)y'+xy=0.$

2) $x(x^2+3x+3)y''+(3x^2+6x+3)y'+(x+1)y=0.$

3) $x(x-1)(x+8)y''+(3x^2-14x-8)y'+(x+2)y=0.$

4) $x(x^2+11x-1)y''+\{3x^2+22x-1)y'+(x+3)y=0.$

Each of the equations 1)-4) is a pull-back of a hypergeometric function by a rational map. In fact, for each of the equations 1)-4) there exist an integral power series $y(x)=\sum_{N=0}^{\infty}c_Nx^N$ satisfying $Ly=0$ and regular at $x=0$.

Apéry's example for $\zeta(2)$ arises from the equation 4) and from a solution of a non-homogeneous equation $Ly=\text{const}\neq 0$ $z(x)=\sum_{N=0}^{\infty}d_Nx^N$ regular at $x=0$ with nearly integral d_N (this is according to the global nilpotence of the corresponding L). Then explicit computations show that c_N/d_N provide with the dense system of approximations to $\zeta(2)$ showing the irrationality of $\zeta(2)$. Similarly, approximations to $\zeta(3)$ arise from the the symmetric square of equation 3).

These examples lead to a method of the construction of sequences of dense approximation to numbers using nearly integral solutions of globally nilpotent equations. Often the corresponding equations are Picard-Fuchs equations satisfied by generating functions of Padé approximants to solutions of special linear differential equations, see examples in [28], [29].

Diophantine approximations suggest the following problem: determine all cases of global nilpotence of Lamé equations.

Our intensive numerical experiments reveal predictable, phenomenon: it seems that, with the exception of equations 1)-4) (and all equations equivalent to them via Möbius transformations), there is *no* Lamé equations with $n=-1/2$ over $\overline{\mathbf{Q}}$ that are globally nilpotent. We put these observations as a

Conjecture2.2. Lamé equations with $n=-1/2$ defined over $\overline{\mathbf{Q}}$ are not globally nilpotent except for 4 classes of equations corresponding to the congruence subgroups, with representatives of each class given by 1)-4).

What are our grounds for this conjecture? First of all, Padé approximation technique allows us to prove one positive result in the direction of this conjecture for the $n=-1/2$ case of the Lamé equation.

Theorem2.3. For fixed $a\in\overline{\mathbf{Q}}$ $(a\neq 0,1)$, there are only finitely many algebraic numbers C of bounded degree d such that the Lamé equation with $n=-1/2$ is globally nilpotent.

Of course, one wants a more specific answer for any other n (e.g. for $n=-1/2$, there are only 4 classes of a and C given above with the global nilpotence). However for half-integral n, there are always $n-1/2$ trivial cases of global nilpotence, where solutions are expressed in terms of elliptic integrals.

We have started the study of globally nilpotent Lamé equations (2.4) or (2.5) with numerical experiments. This ultimately led to Conjecture 2.2. We checked possible equations of the form (2.5) with

$$P(x)=x(x^2-a_1x+a),$$

i.e. 4 singularities at $x = 0, \infty$ and 2 other points, for values of

$$a_1, a \in \mathbf{Z}$$

in the box: $|a|, |a_1| \leq 200$.

For all these equations (2.5) we checked their p-curvature for the first 500 primes. Our results clearly show that with an exception of 4 classes of equations equivalent to 1)-4), any other equation has a large proportion of primes p such that the p-curvature is not nilpotent for any $H \in \mathbf{Q}$!

An interesting p-adic problem arises when, instead of globally nilpotent equations one looks at the nilpotence conditions of p-curvature for a fixed p or, equivalently, when there is a p-integral solution for a fixed p. The only known case (Tate-Deligne), corresponds to Lamé equation with $n = 0$, where the unit root of the ζ-function of the corresponding elliptic curve is expressed in terms of a unique accessory parameter. This example suggests a definition of p-adic spectrum, which we use only for Lamé equations.

We are interested in those $H \bmod p$ for which the p-curvature of (2.5) is nilpotent, and particularly in those p- adic $H \in \mathbf{Z}_p$ for which there exists a solution $y = y(x)$ of (2.5) whose expansion has p-integral coefficients. We call those $H \in \mathbf{Z}_p$, for which such $y(x)$ exists, eigenvalues of (2.5) in the "p-adic domain", and their set we call "an integral p-adic spectrum". The problem of study of arithmetic nature of Lamé equation was proposed by I.M. Gelfand.

To determine p-adic spectrum we conducted intensive symbolic and numeric computations, usint SCRATCHPAD (IBM), MACSYMA (Symbolics) and array processors.

We start with the observations of the "mod p" spectrum as p varies.

I. For noncongruence equations (2.5) with rational $a \neq 0, 1$ (i.e. for an elliptic curve defined over $\mathbf{Q}$ with a point of order 2) there seem always to be infinitely many primes p for which *no* value of the accessory parameter H mod p gives a nilpotent p-curvature (thus mod p spectrum is empty).

Sometimes the first prime p, for which the mod p spectrum of (2.5) is null, occurs quite far. Here are a few statistics for noncongruence equations with rational integers a:

For $a = 3$ the first p's with the null spectrum mod p are: p = 61, 311, 677, 1699, 1783, 1811, 2579, 2659, 3253,... . For $a = 5$ the first p's with the null spectrum mod p are: p = 659, 709, 1109, 1171, 1429, 2539, 2953, 2969, 3019, 3499, 3533, 3803, 3863, 4273, 4493, 4703, 4903, 5279, 5477, 5591, 6011, 7193, 7457, 7583,... . For $a = 4$ the corresponding p's with the null spectrum are: p = 101, 823, 1583, 2003, 3499,... . For $a = 13$ the corresponding list starts at: p = 1451, 1487, 2381,... .

Observation I above was checked by us for all noncongruent $P(x) = x(x^2 - a_1 x + a)$ with integral $a_1 a$ not exceeding 250 in absolute value.

II. An integral p-adic spectrum of equations (2.5) with (p- integral) a has a complicated structure depending on the curve. p-adic spectrum can be null, finite (typically a single element), or infinite, resembling the Cantor set.

Numerical analysis is not easy either. For example, in order to determine the 3-adic spectrum with 14 digits of precision (in the 3-adic expansion), one has to carry out computations with over 2,000,000 decimal digit long numbers!

Example above show how globally nilpotent equations can be used for interesting

arithmetic applications, particularly irrationality proofs. We suggest, as a starting equation, when it is of the second order, an equation corresponding to an arithmetic Fuchsian subgroup. Congruence subgroups of $\Gamma(1)$ and quaternion groups provide with interesting families of globally nilpotent equations.

One can start with equations uniformizing punctured tori with more than one puncture. The complete description of arithmetic Fuchsian groups of signature $(1;e)$ had been provided by Maclachlan and Rosenberger [30] and Takeuchi [31].

For all $(1;e)$ arithmetic subgroups there exists a corresponding Lamé equation with a rational n, uniformized by the corresponding arithmetic subgroup. This way we obtain *70* Lamé equations, all defined over $\mathbf{Q}$. Some of these equations give rise to nearly integral sequences satisfying three-term linear recurrences with coefficients that are quadratic polynomials in n, and have the growth of their denominators and the convergence rate sufficient to prove the irrationality of numbers arising in this situation in a way similar to that of Apéry.

Groups of the signature $(1;e)$ correspond to the Lamé equations

$$P(x)y'' + 1\ over 2P'(x)y' + \{C - \frac{n(n+1)}{4}x\}y = 0$$

with $n + \frac{1}{2} = \frac{1}{2}e$.

In the arithmetic case one looks at totally real solutions of the modified Fricke's identity, which now takes the form:

$$x^2 + y^2 + z^2 - xyz = 2(1 - \cos(\frac{\pi}{e})).$$

Using numerical solution of the (inverse) uniformization problem, we determined the values of the accessory parameters. We display simplest examples:

Here $P(x) = x(x-1)(x-A)$ and:

(1;2)-case:

1) $A = 1/2$, $C = -3/128$,

2) $A = 1/4$, $C = -1/64$;

3) $A = 3/128$, $C = -13/2^{11}$;

4) $A = (2-\sqrt{5})^2$, $C = \sqrt{5}\cdot(2-\sqrt{5})/64$;

5) $A = (2-\sqrt{3})^4$, $C = -(2-\sqrt{3})^2/2^4$;

6) $A = (21\sqrt{33} - 27)/256$,

(1;3)-case

1) $A = 1/2$, $C = -1/36$

2) $A = 32/81$, $C = -31/2^4\cdot 3^4$;

3) $A = 5/32$, $C = -67/2^9\cdot 3^2$;

4) $A = 1/81$, $C = -1/2\cdot 3^4$;

5) $A = (8 - 3\sqrt{7})/2^4$,

3. Continued Fractions and

Applications to Diophantine Approximations.

We start this brief exposition of our results on diophantine approximations with the following translation of our conjectures from §§1-2 on globally nilpotent equation into classical problem of nearly integral solutions to linear recurrences.

Problem. Let u_n be a solution of a linear recurrence of rank r with coefficients that are rational (polynomial) in n:

$$u_{n+r} = \sum_{k=0}^{r-1} A_k(n) \cdot u_{n+k}$$

for $A_k(n) \in \overline{\mathbf{Q}}(n) : k = 0, \ldots, r-1$, and such that u_n are "nearly integral". Then the generating function of u_n is a function whose local expansion represents either an integral of an algebraic function or a period of an algebraic integral, i.e. a solution of Picard-Fuchs-like equation.

For continued fractions this problem can be reformulated.

Problem'. Let us look at an explicit continued fraction expansion with partial fractions being rational functions of indices:

$$\alpha = [a_0; a_1, \ldots, A(n), A(n+1), \ldots],$$

for $A(n) \in \mathbf{Q}(n)$. Let us look then at the approximations P_n/Q_n to α defined by this continued fraction expansion. If the continued fraction representing α is convergent *and* for *some* $\epsilon > 0$

$$|\alpha - \frac{P_n}{Q_n}| < |Q_n|^{-1-\epsilon} :$$

$n \geq n_0(\epsilon)$, i.e. if α is *irrational*, then the sequences P_n and Q_n of numerators and denominators in the approximations to α are arithmetically defined sequences; their generating functions represent solutions of Picard-Fuchs and generalized Picard-Fuchs equations.

How often do such continued fraction expansions do occur, apart from classical cases known to Euler (Hermite in the multidimensional case)? One of the main purposes of our investigation was an attempt to establish, first empirically, that there are only finitely many classes of such continued fraction expansions all of which can be determined explicitly.

In applications to diophantine approximations, a particular attention was devoted to three-term linear recurrences like:

$$n^d \cdot u_n = P_d(n) \cdot u_{n-1} - Q_d(n) \cdot u_{n-2} : n \geq 2$$

for $d \geq 2$. Apart from trivial cases (reducible to generalized hypergeometric functions), our conjectures claim that for every $d \geq 1$, there are only finitely many classes of such recurrences that correspond to deformations of algebraic varieties.

For $d = 2$ (second order equations) we have classified nontrivial three-term recurrences whose solutions are always nearly integral, assuming our integrality conjectures. Most of these recurrences are useless in arithmetic applications. There are a few new ones that give some nontrivial results. Among these recurrences are the following:

i) $2n^2u_n = 2(-15n^2 + 20n - 7) \cdot u_{n-1} + (3n-4)^2 \cdot u_{n-2}$;

ii) $3n^2u_n = (-12n^2 + 18n - 7) \cdot u_{n-1} + (2n-3)^2 \cdot u_{n-2}$;

iii) $n^2u_n = (-12n^2 + 18n - 7) \cdot u_{n-1} + (2n-3)^2 \cdot u_{n-2}$;

iv) $n^2 \cdot u_n = (56n^2 - 70n + 23) \cdot u_{n-1} - (4n-5)^2 \cdot u_{n-2}$.

There is a larger class of rank $r > 2$ linear recurrences of the form

$$n^2 \cdot u_n = \sum_{k=1}^{r} A_k(n) \cdot u_{n-k},$$

all solutions of which are nearly integral. Many of these recurrences give rise to new irrationalities. E.g. we present the following new globally nilpotent equation ($r = 3$):

$$4x(x^3 + 16x^2 + 77x - 2)y'' + 8(2x^3 + 24x^2 + 77x - 1)y' + (9x^2 + 70x + 84)y = 0.$$

Recently, studying Lamé equations we discovered new classes of explicit continued fraction expansions of classical (special) functions, related to problems above. These continued fraction expansions include many Stieltjes-Rogers' continued fractions and are related to elliptic theta-functions [62-63].

Stieltjes-Rogers' expansions (see[33]) include the following examples:

$$\int_0^\infty sn(u,k^2)e^{-ux}du = \frac{1}{z^2 + a-}\frac{1 \cdot 2^2 \cdot 3k^2}{z^2 + 3^2a-}\frac{3 \cdot 4^2 \cdot 5k^2}{z^2 + 5^2a-}\frac{5 \cdot 6^2 \cdot 7k^2}{z^2 + 7a^2 - \ldots} \tag{3.1}$$

$$z\cdot\int_0^\infty sn^2(u,k^2)e^{-uz}du = \frac{2}{z^2 + 2^2a-}\frac{2 \cdot 3^2 \cdot 4k^2}{z^2 + 4^2a-}\frac{4 \cdot 5^2 \cdot 6k^2}{z^2 + 6^2a-}\frac{6 \cdot 7^2 \cdot 8k^2}{z^2 + 8^2a - \ldots}$$

$a = k^2 + 1$.

In the case of expansion (3.1) the approximations P_m/Q_m to the integral in the left hand side of are determined from a three-term linear recurrence satisfied by P_m and Q_m

$$(2m+1)(2m+2)\phi_{m+1}(z) = (z + (2m+1)^2a)\phi_m(z) - 2m(2m+1)k^2\phi_{m-1}(z).$$

Here $\phi_m = P_m$ or Q_m, and Q_m are orthogonal polynomials. The generating function of Q_n satisfy a Lamé equation in the algebraic form with a parameter $n = 0$.

These special continued fraction expansions can be generalized to continued fraction expansions associated with any Lamé equation with an arbitrary parameter n.

For $n = 0$ these closed form expressions represent the Stieltjes-Rogers expansions. For $n = 1$ two classes of continued fractions from [27] have arithmetic applications, because for three values of the accessory parameter H (corresponding to e_i-nontrivial 2nd order points) the Lamé equation is a globally nilpotent one and we have p-adic as well as archimedian convergence of continued fraction expansions. This way we obtain the irrationality and

bounds on the measure of irrationality of some values of complete elliptic integrals of the third kind, expressed through traces of the Floquet matrices. Similarly, for an arbitrary integral $n \geq 1$, among continued fraction expansions, expressed as integrals of elliptic θ-functions, there are $2n+1$ cases of global nilpotence, when continued fractions have arithmetic sense and orthgonal polynomials have nearly integral coefficients.

Among new explicit continued fraction expansions is the expansion of the following function generalizing Stieltjes-Rogers:

$$\int_0^\infty \frac{\sigma(u-u_0)}{\sigma(u)\sigma(u_0)} e^{\zeta(u_0)u} du.$$

The three-term linear recurrence determining the J-fraction for the corresponding orthogonal polynomials has the following form:

$$Q_N(x) = Q_{N-1}(x)\cdot\{(l+k^2)\cdot(N-1)^2+x\}+Q_{N-2}(x)\cdot k^4\cdot(N-1)^2\cdot N\cdot(N-\frac{1}{2})\cdot(N-\frac{3}{2})\cdot(N-\frac{5}{2}).$$

Here $x = sn^2(u_0;k^2)$.

The more general J-fraction of the form

$$\ddots\, b_{n-1} + x - \cfrac{a_{n-1}}{b_n + x - \cfrac{a_n}{b_{n+1} + x - \ddots}},$$

with

$$a_n = k^4 \cdot n(n+1)\cdot(n+\frac{1}{2})(n-\frac{1}{2})\cdot\{(n-1)\cdot(n-\frac{1}{2}) - \frac{m\cdot(m+1)}{4}\};$$

$$b_n = (1+k^2)\cdot(n-1)^2 : n \geq 2$$

is convergent to the integral of the form

$$\int_0^\infty \sqcap_{i=1}^m \frac{H(u-u_i)}{\Theta(u)\theta(u_i)} e^{-Z(u_i)u} du.$$

The generating function of the corresponding orthogonal polynomials is expressed in terms of solutions of a Lamé equation with parameter $m \geq 1$.

Arithmetic applications, particularly to the determination of measures of irrationality, of classical constants often require E- or G-function representations. Rapidly convergent nearly integral power series expansions are the most efficient way to construct Padé approximations and determine the arithmetic nature of classical constants. We have developed the theory of new identities giving such representations [32]. They generalize Ramanujan's quasiperiod relations [34] that give generalized hypergeometri series of multiple of l/π. To introduce Ramanujan's series we first need Eisenstein's series:

$$E_k(\tau) = 1 - \frac{2k}{B_k}\cdot\sum_{n=1}^{\infty}\sigma_{k-1}(n)\cdot q^n$$

for $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$, and $q = e^{2\pi i\tau}$. In the $E_k(\tau)$ notations, the quasiperiod relation is expressed by means of the function

$$s_2(\tau) \stackrel{\text{def}}{=} \frac{E_4(\tau)}{E_6(\tau)} \cdot (E_2(\tau) - \frac{3}{\pi \text{Im}(\tau)}), \tag{3.2}$$

which is nonholomorphic but invariant under the action of $\Gamma(1)$.

Ramanujan proved in [34] that this function admits algebraic values whenever τ is imaginary quadratic.

The Clausen identity gives the following ${}_3F_2$- representation for an algebraic multiple of $1/\pi$, following from (3.2):

$$\sum_{n=0}^{\infty} \{\frac{1}{6}(1 - s_2(\tau)) + n\} \cdot \frac{(6n)!}{(3n)!n!^3} \cdot \frac{1}{J(\tau)^n} \tag{3.3}$$

$$= \frac{(-J(\tau))^{1/2}}{\pi} \cdot \frac{1}{2(d(1728 - J(\tau))^{1/2}}$$

Here $\tau = (1 + \sqrt{-d})/2$. If $h(-d) = 1$, then the second factor in the right hand side is a rational number. The largest one class discriminant $-d = -163$ gives the most rapidly convergent series among those series where all numbers in the left side are *rational*:

$$\sum_{n=0}^{\infty} \{c_1 + n\} \cdot \frac{(6n)!}{(3n)!n!^3} \frac{(-1)^n}{(640,320)^n} = \frac{(640,320)^{3/2}}{163 \cdot 8 \cdot 27 \cdot 7 \cdot 11 \cdot 19 \cdot 127} \cdot \frac{1}{\pi}.$$

Here

$$c_1 = \frac{13,591,409}{163 \cdot 2 \cdot 9 \cdot 7 \cdot 11 \cdot 19 \cdot 127}$$

and $J(\frac{1+\sqrt{-163}}{2}) = -(640,320)^3$.

Ramanujan provides instead of this a variety of other formulas connected mainly with the three other triangle groups commensurable with $\Gamma(1)$. All four classes of ${}_3F_2$ representations of algebraic multiples of $1/\pi$ correspond to four ${}_3F_2$ hypergeometric functions (that are squares of ${}_2F_1$-representations of complete elliptic integrals via the Clausen identity). These are

$${}_3F_2(\frac{1}{2}, \frac{1}{6}, \frac{5}{6}; 1, 1; x) = \sum_{n=0}^{\infty} \frac{(6n)!}{(3n)!n!^3} (\frac{x}{12^3})^n$$

$${}_3F_2(\frac{1}{4}, \frac{3}{4}, \frac{1}{2}; 1, 1; x) = \sum_{n=0}^{\infty} \frac{(4n)!}{n!^4} (\frac{x}{4^4})^n$$

$${}_3F_2(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, 1; x) = \sum_{n=0}^{\infty} \frac{(2n)!^3}{n!^6} (\frac{x}{2^6})^n$$

$${}_3F_2(\frac{1}{3}, \frac{2}{3}, \frac{1}{2}; 1, 1; x) = \sum_{n=0}^{\infty} \frac{(3n)!}{n!^3} \cdot \frac{(2n)!}{n!^2} (\frac{x}{3^3 \cdot 2^2})^n.$$

Representations similar to (3.3) can be derived for any of these series for any singular moduli $\tau \in \mathbf{Q}(\sqrt{-d})$ and for any class number $h(-d)$, thus extending Ramanujan list ad infinum. For a simple recipe to generate these new identities, see [32].

Logarithms of quadratic units, like π, can be represented as values of convergent series satisfying Fuchsian linear differential equations. This holds for $\log \epsilon_{\sqrt{k}}$ of a fundamental unit $\epsilon_{\sqrt{k}}$ of a real quadratic field $\mathbf{Q}(\sqrt{k})$. To represent this number as a convergent series (in, say, $1/J(\tau)$) one uses Kronecker's limit formula expressing this logarithm $\log \epsilon_{\sqrt{k}}$ in terms of products of values of Dedekind 's Δ-function ("Jugendtraum").

Ramanujan's and other similar identities that express π and other similar numbers in terms of values of G-functions very close to the origin, give us the basis of applications of our Padé approximation techniques to these G-functions. In such applications the exponent in the measure of diophantine approximations strongly depends on the proximity of an evaluation point to the point of expansion of a G-function. To make this dependence explicit, we quote the following our result [14]:

*Theorem*3.1[14]. Let $f_1(x), \ldots, f_n(x)$ be G-functions satisfying linear differential equations over $\mathbf{Q}(x)$. Let $r = a/b$, with integers a and b, be very close to the origin. Then we get the following lower bound for linear forms in $f_1(r), \ldots, f_n(r)$.

For arbitrary non-zero rational integers $H_1, \ldots, H_n$ and $H = \max\{|H_1|, \ldots, |H_n|\}$, if $H_1 f_1(r) + \ldots + H_n f_n(r) \neq 0$,

$$|H_1 f_1(r) + \ldots + H_n f_n(r)| > |H_1 \cdots H_n|^{-1} \cdot H^{1-\epsilon}$$

provided that r is very close to 0:

$$|b| \geq c_1 \cdot |a|^{n(n-1+\epsilon)}$$

and $H \geq c_2$ with effective constants $c_i = c_i(f_1, \ldots, f_n, r, \epsilon)$. If r is not *as* close to 0, we get only

$$|H_1 f_1(r) + \ldots + H_n f_n(r)| > H^{\lambda-\epsilon}$$

for $\lambda = -(n-1)\log|b|/\log|b/a^n| (< 0)$.

We will present some of the results for numbers connected with π based on effective Padé approximations with schemes described in [29], [35]. The first bound is connected with Ramanujan-like series:

$$|\pi\sqrt{2} - p/q| > |q|^{-16.67\ldots}$$

for rational integers $p, q : |q| \geq q_0$.

For $\pi\sqrt{3}$ we use different system of Padé-type approximations [28]. Below we present the corresponding integral representations, leading to the bound:

$$|\pi\sqrt{3} - p/q| > |q|^{-5.791\ldots}$$

for $|q| \geq q_2$.

To cover large classes of numbers we generalize the Ramanujan theory quasi-period relations to the general CM-varieties. Particularly interesting applications arise for arithmetic triange (quaternion groups [32]). To present some examples of such relations, we

look at the automorphic function $\emptyset(\tau)$ for the arithmetic triangle group Γ, normalized by its values at vertices.

An analog of $s_2(\tau)$ in (3.2) that is a nonholomorphic automorphic form for Γ is

$$-\frac{1}{\phi'(\tau)} \cdot \{\frac{\phi''(\tau)}{\phi'(\tau)} - \frac{i}{\mathrm{Im}(\tau)}\}.$$

For $\phi(\tau) = J(\tau)$ one gets $s_2(\tau)$.

For example, let us look at a quaternion triangle group $(0;3;2,6,6)$. In this case, instead of an elliptic Schwarz formula one has the following representation of the normalized automorphic function $\phi = \phi(\tau)$ in H in terms of hypergeometric functions:

$$\frac{\tau + i(\sqrt{2}+\sqrt{3})}{\tau - i(\sqrt{2}+\sqrt{3}} = -\frac{3^{1/2}}{2^2 \cdot 2^{1/6}} \cdot \{\frac{\Gamma(1/3)}{\sqrt{\pi}}\}^6 \cdot \frac{F(\frac{1}{12}, \frac{1}{4}; \frac{5}{6}; \phi)}{\phi^{1/6} \cdot F(\frac{1}{4}, \frac{5}{12}; \frac{7}{6}; \phi)}.$$

The role of π in Ramanujan's period relations is occupied in
$(0,3;2,6,6)$-case by the transcendence $\{\frac{\Gamma(1/3)}{\pi}\}^6$.

Thus, generalizations of Ramanujan identities allow us to express constants, such as π and other Γ-factors, as values of rapidly convergent series with nearly integral coefficients in a variety of ways, with convergence improving as the discriminant of the corresponding singular moduli increases.

Another interesting dimension in all these identities is an ability to add p-adic interpretation to them. We are developing now a variety of p-adic analogs of such identities to be used for diophantine and for numerical applications.

We refer also to our papers [43], [44] and [Maine] for a variety of applications of effective Padé approximation techniques to measures of diophantine approximations. These techniques, particularly from [29] and [37] are based on multiple integrals of Pochhammer type and the combination of symbolic and numerical methods to determine the best sequences of dense approximations to such numbers as π, ln 2, $\sqrt{[3]2}$.

4.Computer Algebra and Numerical Computations.

Our work on computer algebra systems included the development of new algorithms of symbolic manipulation with algebraic numbers and functions, solution of algebraic, differential and integral equations, and development of algorithms combining computer algebra and numerical methods. We refer to [27] for description of computer algebra methods arising from elliptic curves and Abelian varieties with applications to number theory, including diophantine approximations, primality testing and factorization, to algebraic topology, K - theory and superstring interpretation, and to uniformization theory.

Some of our methods developed for the study of arithmetic properties of linear differential equations, were used in the development of computer algebra systems themselves. This includes new methods of power series manipulation, and, in particular, fast algorithms of power series expansions of solutions of differential equations [31], [61]. These methods also led to new algorithms of explicit integration of algebraic functions, including the determination of all cases, when this is possible (this is closely related to our solution of the Grothendieck problem, see [21], [27]). In the environment of computer algebra systems we were developing some new number-theoretic algorithms, including new algorithms of bignum operations, factorization, and new modular algorithms of fast convolution. All this work was conducted in the environment of traditional computer algebra systems, particularly within MACSYMA and SCRATCHPAD II environments. Of these we mainly benefited from the impressive capabilities of SCRATCHPAD (IBM), which showed itself capable of handling very large computational tasks and providing with a very flexible programming support. Our involvement with SCRATCHPAD benefited our theoretical and applied work, and gave us high hopes for its future utilization, especially when it will be widely released.

At the same time as we used the general purpose computer algebra systems, we are developing specialized packages for specific tasks and applications. An acute need in such specialized packages is felt in the development of software support for handling large computational tasks for vector and parallel machines based on modeling or simulation of realistic physical or engineering problems. Among classes of problems that we are interested in, we would like to mention: solution of boundary value problems, study of multi-dimensional elliptic problems for domains of complex boundary structure, development of parallel algorithms for solution of two-dimensional and three-dimensional aero- and hydro- dynamic problems, and the study of astrophysical models, combining chemistry, thermodynamics, gravitation and relativistic effects. Development of vectorized and parallelized codes for these problems is necessary because large realistic modelling is possible only on supercomputers and massively parallel machines, with GigaFlops of performance and GigaBytes of storage. The issue of parallelization and, particularly, intercommunication between many processors is an important and challenging mathematical problem in itself. It leads to a variety of outstanding questions of complexity, number theory and graph theory. We present in the Appendix results of one of our works (jointly with M. M. Denneau) on the construction of optimal sparse networks.

We use computer algebra systems to develop optimized numerical algorithms for solu-

tion of specific equations and modelling of physical processes. We also use computer algebra systems to map developed algorithms to vector or parallel organizations of large machines and supercomputers. Both of these tasks are so involved and time consuming that without support of the computer algebra environment they cannot be efficiently accomplished. As an example of our involvement, let us describe our work on optimization of two -and three- dimensional aero -and hydro- dynamic codes. Part of this work started a few years ago, when we together with K. Prendergast (Dept. of Astrophysics, Columbia University), conducted numerical simulation of underground explosion in cavities of varying geometry for the Hudson Institute and DOD. The code was the so-called beam-scheme, quite robust general hydrodynamics code used in many problems of fluid and flow dynamics.

Recently we returned to fluid and flow problems, when we started to look on examples of large sale problems that cannot be solved without parallel processing, due to their numerical complexity and memory requirements. One of the best examples of such modeling is provided by astrophysical and cosmological numerical experiments. In these experiments the goal is to simulate all known complexities of galaxy formation and evolution. This includes interactions of stars, interstellar matter, gravitational fields, chemistry and thermodynamics and the presence of dark matter. Three - dimensional simulations, involving large number of objects (requiring at least a million of grid points) over a long period of galaxy life are some of the most numerically and memory intensive challenging tasks. Optimization of numerical algorithms and their parallelization are necessary in order to tackle these problems and give to astrophysicists sufficient amount of data to study, and to develop better models of physical processes.

Our current work involves three - dimensional multifluid models, combining chemistry, thermodynamics, gravitational potential and global gravitational effects, and a detailed look at each cell. In development of better numerical methods computer algebra techniques and tools were an important component. The code itself is extremely floating point intensive, because it involves at every discrete time step computations of a variety of elementary and special functions, including EXP, ERFC (error functions) and several complicated integrals. For example, per grid point (a million of them), per discrete step (about 10^5 of them), one has to compute with *high accuarcy* values of 4 integrals of the form

$$\frac{\int_a^\infty e^{-(x-d)^2} x^k \, sqrt x^2 + c^2}{sqrt x^2 + e^2 \, dx.}$$

Standard numerical quadrature methods (e.g. Gauss - Hermite) and library calls (sach as NAG, IMSL, CRAY, IBM ESSL-V, etc.) make this part of computations prohibitively expensive. Using our methods of fast computation of solutions of linear differential equations we developed with help of computer algebra systems a fast method of high precision evaluation of this and other similar integrals. Such a development, combined with the full parallelization now allows to implement three - dimensional code on multiprocessor machines, and make the runtime of the code tolerable for modeling of a variety of initial and boundary conditions. We are now running a variety of new codes for these astrophysical problems in two fluid version for systems with spherical and cylindrical symmetries on a variety of general purpose and special computers (some of the hardware designed by us). We are preparing now for large runs of three-dimensional codes on supercomputers.

5. Basic Algebraic Complexities.

Algebraic complexities were originally developed to describe the number of operations: multiplications (multiplicative complexity) and additions (additive complexity) in traditional algebraic structures: rings of matrices and rings of polynomials. While the operation of addition in such rings is straightforward, it is by no means obvious how to realize the operation of multiplication in these rings in the least number of operations on primitives (numbers), particularly in the least number of multiplications.

The algebraic complexity problem which is the richest in its underlying structure is the problem of fast polynomial multiplication. Among other problems reducible to that one can mention: fast multiplications of multiple-precision numbers, gdc's in polynomial rings, Hankel matrix multiplication, computation of rational and Padé approximations, computation of finite Fourier transformations,...etc. Significant progress in this problem, due to Winograd, Fiduccia, Strassen and others, was concentrated mainly on minimal multiplicative complexities of polynomial multiplication over infinite *fields*. For practical, particularly hardware implementations, and in applications to fast multiplication of multiple precision numbers all algorithms have to be considered over a ring $\mathbf{Z}$ or $\mathbf{Z}[1/2]$.

We have developed new low complexity algorithms arising from interpolation on algebraic curves of positive genus and on arbitrary algebraic and arithmetic surface. We precede the description of these algorithms with a short exposition of complexity count and Toom-Cook ordinary interpolation method.

The (multiplicative) complexity is well defined for systems of bilinear forms [38-39]. Let $\mathbf{A}$ be a ring (of scalars), and suppose given s bilinear forms in variables $\overline{x} = (x_1, \ldots, x_m)$ and $\overline{y} = (y_1, \ldots, y_n)$ with coefficients from $\mathbf{A}$:

$$z_k = \sum_{i=1}^{m} \sum_{j=1}^{n} t_{i,j,k} x_i y_j : k = 1, \ldots, s. \tag{5.1}$$

One of the most widely used definitions of the multiplicative complexity of computation of a system (5.1) over $\mathbf{A}$ is that of the range of the $m \times n \times s$ tensor $(t_{i,j,k})$ over $\mathbf{A}$ (Strassen). A nonzero tensor T is said to be of rank 1 over $\mathbf{A}$. If there are three $\mathbf{A}$-vectors $(a_1, \ldots, a_m), (b_1, \ldots, b_n), (c_1, \ldots, c_s)$ such that

$$t_{i,j,k} = a_i \cdot b_j \cdot c_k \text{ for all } i, j, k.$$

Then, the rank of a tensor T over $\mathbf{A}$ is the minimal number, $\delta_{\mathbf{A}}$, such that T is expressible as a sum of $\delta_{\mathbf{A}}$ rank 1 tensors over $\mathbf{A}$. This scheme of evaluation of (5.1) is called a bilinear scheme (normal or noncommutative). In this scheme for the evaluation of a system of bilinear forms (5.1) one forms δ products of linear combinations of $x's$ and $y's$:

$$w_l = (a_{l1} x_1 + \ldots + a_{lm} x_m)(b_{l1} y_1 + \ldots + b_{ln} y_n), \tag{5.2a}$$

$l = 1, \ldots, \delta$; and then obtains z's in (5.1) as linear combinations of these products:

$$z_k = c_{k1}w_1 + \ldots + c_{k\delta}w_\delta : k = 1, \ldots, s. \tag{5.2b}$$

In this definition of multiplicative complexity one counts only *non-scalar* (essential) multiplications, i.e. δ multiplications in (5.2a).

For future reference, we express the bilinear algorithm (5.2a)- (5.2b) of computation of bilinear forms (5.1) in the following algebraic form

$$\overline{z} = C \cdot (A\overline{x} \otimes B\overline{y}) \tag{5.3}$$

for matrices $C = (c_{kl})(\in M_{s\times\delta}(\mathbf{A}))$, $A = (a_{li})(\in M_{\delta\times m}(\mathbf{A}))$, $B = (b_{lj})(\in M_{\delta\times n}(\mathbf{A}))$.

If one extends the ring of scalars by inverting primes and by adding algebraic numbers (most notable roots of unity), one can often significantly reduce the multiplicative complexity of computation of systems of bilinear forms. This is the case of polynomial multiplications. Unfortunately, simultaneously the total computational complexity of the algorithms (as reflected in the total number of bit operations) becomes unbearably high. That is why the "least complex" algorithms of polynomial multiplication, developed by Toom-Cook-Winograd [38-39] are unattractive in practice for (relatively) large degrees. We describe briefly these algorithms because they provide clues to the use of interpolation technique.

The basic idea underlying the Toom-Cook method is to use interpolation is reconstruction of coefficients of a product of two polynomials. Namely, let us consider the bilinear problem (5.1) corresponding to the multiplication of two polynomials of degree $m-1$ and $n-1$, respectively:

$$P(t) = \sum_{i=1}^{m} x_i t^{i-1},\ Q(t) = \sum_{j=1}^{n} y_j t^{j-1}. \tag{5.4}$$

One is interested in the efficient determination of coefficients z_k of a polynomial product:

$$R(t) \stackrel{\text{def}}{=} P(t) \cdot Q(t),\ R(t) = \sum_{k=1}^{m+n-1} z_k t^{k-1}. \tag{5.5}$$

A "high school" scheme of evaluation of (5.5) takes mn multiplications (but no scalar multiplication). Toom noticed, however, that one can identify $R(t)$, if one knows its values at $m+n-1$ distinct points $a_k : k = 1, \ldots, m+n-1$ in terms of the Lagrange interpolation formula:

$$R(t) = \sum_{k=1}^{m+n-1} R(\alpha_k) \cdot \frac{\sqcap_{l\neq k}(t-\alpha_l)}{\sqcap_{l\neq k}(\alpha_k - \alpha_l)}. \tag{5.6}$$

Thus if one assumes the ring $\mathbf{A}$ to *include* α_i and $(\alpha_i - \alpha_j)^{-1}$ (for $i \neq j$), then one can reconstruct all $m+n-1$ coefficients z_k of $R(t)$ from $R(\alpha_k)$ using scalar multiplications only. To compute $R(\alpha_k)$ one needs then only $m+n-1$ nonscalar multiplications:

$$R(\alpha_k) = P(\alpha_k) \cdot Q(\alpha_k), \tag{5.7}$$

where $P(\alpha_k)$ and $Q(\alpha_k)$ are computed by the Horner scheme in (5.4) with scalar (from $\mathbf{A}$) multiplications only.

The Toom-Cook scheme was generalized by Winograd, who showed that one can evaluate $R(t)$ mod arbitrary relatively prime polynomials $D_\alpha(t)$, and then bring the results together using the Chinese remainder theorem and without any *nonscalar* multiplications to obtain $R(t)$ mod $\sqcap_\alpha\ D_\alpha(t)$. Toom-Cook took $D_\alpha(t) = t - \alpha$, while Winograd considered $D_\alpha(t)$ as cyclotomic polynomials. Optimal results of polynomial multiplication and multiplications in algebras $\mathbf{A}[x]/(T(x))$ (i.e. polynomial multiplication mod$T(x)$), when $\mathbf{A} = K$ is a *infinite field*, are described by the following Winograd's

*Theorem*5.1: Let K be an infinite field. Then the multiplicative complexity, $\delta_K(m,n)$, of multiplication of polynomials of degrees $m-1$ and $n-1$ over $K[x]$ is $m+n-1$ exactly. The multiplicative complexity, $\delta_K(T)$, of multiplication of polynomials in $K[x]$ mod $T(x)$ for $T(x) \in K[x]$ is equal to $2n-k$, where k is a number of distinct irreducible factors of $T(x)$ in $K[x]$.

Though these results are strong, none of them can be applied in practice (starting from $m \geq 4$, $n \geq 4$) because in any of the schemes of minimal multiplicative complexity (over, say, $\mathbf{Q}$), the scalar multiplications not counted as "actual multiplications"–involve scalars that are too large to be bounded as just a few more additions. Also one has to divide by large integers having no advantageous binary structure.

Not only the large sizes of scalars make the number of additions prohibitively - exponentially - large, but the coefficients become large integers requiring fast bignum multiplications. The total number of operations, counted in terms of single precision additions and multiplications, in the "minimal multiplicative complexity" algorithms of Theorem 5.1 significantly exceeds the number of operations in the standard high school methods of polynomial multiplication. This makes the minimal multiplication complexity algorithms of Theorem 5.1 unsuitable for practical implementation. The minimal multiplication routines of Theorem 5.1 found their place, though, in Winograd's short prime length DFT algorithms; they are well suited for iterations when one uses them only for short data sizes.

It is more efficient in practice to use fast multiplication routines with as little division by scalars as possible (and no prime inversions!). From the point of view of hardware realization it is preferable to have division by power of 2 only, i.e. one should consider minimal multiplication schemes over $\mathbf{A} = \mathbf{Z}[1/2]$. It was realized some time ago ([Schonhage–Strassen] [41-42]) that one can achieve fast multiplication of polynomials with division by 2 only, if one considers polynomial multiplications modulo cyclotomic polynomials $(x^{2^n}-1)$. This general scheme, coupled with fast Fourier transforms in finite fields (modulo factors of Fermat numbers) was used first by [41] to achieve asymptotically fast multiplication of large integers. In that method one can multiply two n-bit integers in time $0(n \log n \log\log n)$ (i.e. in the many bit-operations).

We see that linear upper bounds for multiplicative complexities of polynomial multiplication (given, say in Theorem 5.1) imply subexponential lower bounds for additive complexities of these algorithms. The best upper bounds for the multiplicative complexity of computation of polynomial multiplication via certain versions of FFT are $0(n \log n)$ for polynomials of degrees bounded by n, even though divisions by powers of 2 (shifts) are necessary in this scheme. No nonlinear lower bound is known for algebraic complexities in

this problem, and the best total complexity known is $0(n \log n \log\log n)$.

To see what can be the **Z**-algorithms of fast polynomial multiplication, we have to consider first the reductions mod p, and to look at algebraic complexities of polynomial multiplications of $\mathbf{A} = \mathbf{F}_p$, particularly over $\mathbf{A} = \mathbf{F}_2$.

Over finite fields there is *no* simple answer to the minimal multiplicative complexity of polynomial multiplication like one given in Theorem 5.1 for infinite fields. If $\delta_{\mathbf{A}}(m,n)$ denotes the minimal multiplicative complexity of multiplication of polynomials of degrees $m-1$ and $n-1$, respectively, with a ring of constants $\mathbf{A}$, then we always have $\delta_K(m,n) \geq m+n-1$ for an arbitrary field K of scalars, but this inequality becomes equality *only* when K has at least $m+n-2$ elements.

As it turns out, lower bounds for multiplicative complexities over finite fields, and, as a consequence, over the ring $\mathbf{A}$ of scalars, are much stronger than the one given above. To obtain them, though, the algebraic theory of linear codes has to be introduced.

Let us recall the basics of linear codes. In the theory of linear codes one considers vector spaces $\mathbf{F}_q^n$ of dimension n over a finite field $\mathbf{F}_q$. A linear subspace is called a linear code. A linear code is the null space of a parity check matrix of the code, and a basis of the code form the rows of the matrix, called a generator matrix. In addition to n, two more parameters: k and d are associated with a code (called an $[n,k,d]$-code). First, we denote by k the dimension of the code over $\mathbf{F}_q$. Second, by the weight of the code, denoted d, we understand the minimal number of nonzero coordinates of all nonzero vectors from the code with respect to a fixed basis of the vector space.

We will demonstrate the relationship between (multiplicative) complexity of multiplication in k-algebras and the Hamming problem for linear codes over k in the most general situation:

*Corollary*5.2 : If $\mathbf{A}$ is a $\mathbf{F}_q$- algebra of dimension n over $\mathbf{F}_q$ and without zero divisors, then every realization of multiplication in $\mathbf{A}$ over $\mathbf{F}_q$ as a bilinear algorithm with $\delta = \delta_{\mathbf{F}_q}(\mathbf{A})$ nonscalar multiplications over $\mathbf{F}_q$ gives rise to a $[\delta,n,n]$-linear code over $\mathbf{F}_q$.

Corollary 5.2 includes, in particular, all finite extensions of prime fields.

The proof of Corollary 5.2 also provides important clues as to matrices A, B and C in the algebraic form (5.3) of the algorithm of multiplication in k-algebra $\mathbf{A}$.

Using Corollary 5.2 and known lower bounds from the theory of linear codes, one can bound from below the multiplicative complexity of polynomial multiplication. One of the best linear code bounds is (MRRW) proved in [43]. According to this bound for $q=2$, if $n \to \infty$ and there is a sequence of $[n,k,d]$-codes with $d/n \to \delta$, then $R \stackrel{\text{def}}{=} k/n \leq H_2(1/2 - \sqrt{\delta - \delta_2})$. Here the entropy function is $H_q(x) = -x\log_q x - (1-x)\log_q(1-x)$. Such bound was used in [44] to bound the multiplicative complexity $\delta_{\mathbf{F}_2}(n,m)$ of multiplication over $\mathbf{F}_2$ of polynomials of degrees $n-1$ and $m-1$ respectively: $\delta_{\mathbf{F}_2}(n,n) \geq 3.52 \cdot n$, and $\delta_{\mathbf{Z}}(n,n) \geq 3.52 \cdot n$ for a sufficiently large n.

Our new results show that similar lower bounds hold for multiplicative complexity of multiplication in finite extensions of $\mathbf{F}_q$. Indeed let $K = \mathbf{F}_q[t]/p(t)$, for an irreducible polynomial $p(t)$ of degree n in $\mathbf{F}_q[t]$, so $K \cong F_{q^n}$; and let $\delta_{\mathbf{F}_q}(K)$ denote the minimal multiplicative complexity in the field K over $\mathbf{F}_q$. Then Corollary 5.2 implies that there exists a $[\delta_{\mathbf{F}_q}(K),n,n]$-linear code over $\mathbf{F}_q$. Thus we deduce the lower bound

$$\delta_{\mathbf{F}_2}(K) \geq 3.52n \tag{5.8}$$

for sufficiently large $n = [K : \mathbf{F}_2]$.

For $q \geq 2$ one gets less sharp bounds from known bounds on optimal codes. For example, the Plotkin bound implies

$$\delta_{\mathbf{F}_q}(K) \geq (2 + \frac{1}{q-1}) \cdot n$$

as n $[K : \mathbf{F}_q] \to \infty$.

We found an important phenomenon for polynomial multiplication over finite fields drastically different from the infinite field case. In the case of an infinite field of constants k, Winograd's Theorem 5.1 shows that multiplication mod an irreducible polynomial $p(t)$ of degree n takes as many essential multiplications as that of multiplication of two polynomials of degree $n-1$. It is *no* longer true in the finite field case, when we always have $\delta_k(k[t]/(p(t))) + C \cdot n \leq \delta_k(n,n)$ for a positive constant $C(= C(k))$ and an arbitrary polynomial $p(t)$ of degree n over a finite field k. We conjecture, that in fact

$$\lim_{n \to \infty, [K:k]=n} \frac{\delta_k(n,n)}{\delta_k(K)} = 2.0.$$

The *lower* bounds for multiplicative complexities of polynomial multiplication are always linear in n, and seem far from the best nonlinear upper bound $0(n \log n)$ for finite field polynomial multiplication. We have shown that the upper bound can be brought down to a linear one with the constant comparable to that of the lower bound.

To describe new algorithms, we look at arbitrary meromorphic (rational) functions. These rational functions, like polynomial can be represented in a variety of ways. They can be represented as $P(t)/Q(t)$, in terms of its residues: $c_0 + \sum_{i=1}^{n} c_i/(t - \alpha_i)$, or by its values at a given set of points. The last representation is the interpolation. Interpolation formulas are the basis of the Toom- Cook algorithm of the fast polynomial multiplication. Apparently the interpolation algorithm always has the lowers multiplicative complexity. The most important polynomial multiplication problem is that of multiplication in a finite extension $K = k[t]/((p(t))$ for a fixed $p(t) \in k[t]$ and field of constants k. Let $\delta_k(K)$ be the multiplicative complexity of multiplication in K over k. For an infinite k and K of degree n over k, $\delta_k(K) = 2n - 1$ by Theorem 5.1, and all algorithms realizing this multiplicative complexity are interpolation algorithms, interpolating products $x(t) \cdot y(t) \mathrm{mod} p(t)$ at $2n-1$ distinct points of $k\mathbf{P}^1$. What happens if k is finite, say $\mathbf{F}_2$? There are not enough points interpolate at. What usually is done, is an extension of the field of constants (till there are enough points for interpolation) and the nested rule for multiplication in the composition extension of fields. This is expressed by a trivial "multiplication rules":

$$\mu_k(K) \leq \mu_k(\mathcal{L}) \cdot \mu_{\mathcal{L}}(K) \tag{5.9}$$

where $k \subset \mathcal{L} \subset K$. The rules (5.9) are the basis of all previously known and currently used fast convolution algorithms (including Nussbaumer's and Schonhage- Strassen technique).

We found a novel way of interpolation by means of representing the set of points where one interpolates as a divisor on an algebraic curve. To represent this interpolation

in a more abstract way one has to use the standard language of places from the theory of algebraic functions of one variable. See [45-46].

By a place of a field K we understand an isomorphism $\phi : K \to \Sigma \cup \{\infty\}$, where Σ is a field and $\phi(a) = \infty$, $\phi(b) \neq 0, \infty$ for $a, b \in K$. There is an obvious correspondence between places and valuations (which we will use). All places on an algebraic function are extensions of places from corresponding rational function field, and places of a rational function field $K = k(x)$ over a field d correspond either to (i) irreducible polynomials $p(x)$ in $k[x]$, or to (ii) x^{-1}. We deal only with algebraic function fields in one variable. Any such field K over the field of constants k can be represented in the form $K = k(x, y)$, where x is (any) transcendental element of K over k, and $l, \ldots, y^{d-1}$ is the basis of K over $k(x)$, $[K : k(x)] = d$. For an arbitrary place P of K let k_P be a field such that P is an isomorphism onto $k_P \cup \{\infty\}$. We denote by v_P the normed valuation with values in Z, corresponding to P. The degree f_P of k_P over k is called the degree of P. A divisor of K is an element of the free Abelian group generated by the set of places of K. The places themselves are called prime divisors. We write divisors additively: $A = \Sigma_P v_P(A) \cdot P$, where $v_P(A)$ are integers among which only finitely many are nonzero. A divisor A is called an integral one if $v_P(A) \geq 0$ for every P. A divisor A divides B, if $B - A$ is integral. The degree $d(A)$ of a divisor A is an integer $d(A) = \Sigma_P f_P v_P(A)$. With every element $X \in K$ one associates a principal division $(X) = \Sigma_P v_P(x) \cdot P$.

The most important object is the vector space $L(-A)$ over k consisting of all functions $X \in K$ such that the divisor $(X) + A$ is positive. If g denotes the genus of K, then the dimension of $L(-A)$ over k is determined by the Riemann-Roch theorem. To formulate it, we denote by C an equivalence class of divisors in K (modulo the principal ones- (X) for $X \in K$), and by $N(C)$ the dimension of C, i.e. the maximal number of linearly independent integral divisors in this class. Then $N(C)$ is equal to $\dim L(-A)$ for any A from C. The Riemann-Roch theorem states that

$$N(C) = d(C) - g + 1 + i(C), \tag{5.10}$$

where $i(C) = 0$ is the index of speciality of C, $i(C) = N(W - C)$, where W is the class of all differentials on K (canonical class). In particular, $N(C) = d(C) - g + 1$ if $d(C) > 2g - 2$ or $d(C) = 2g - 2$ and $C \neq W$.

In applications. k is often a finite field of characteristic $p > 0$ with $q = p^m$ elements. If P is any prime divisor of K–an algebraic function field with the field of constants k–then the number of elements in the residue field k_P is called the norm of P and is denoted as: $N(P) = q^{d(P)}$. This definition is extended to all divisors: $N(A) = q^{d(A)}$. With an algebraic function field K/k one can associate a ζ- function:

$$\zeta(K; s) = \sum_A (N(A))^{-s}, s > 1$$

where A runs over all integral divisors of K/k.

The interpolation technique on a curve corresponding to K/k proceeds as follows. To represent a finite extension k_P of k of degree $d(P)$ one looks for a positive divisor B such that the natural mapping

$$P : L(-B) \to k_P$$

is onto. This way we find a basis of $k_{\mathcal{P}}$ over k among element of $L(-\mathcal{B})$. E.g. if $K = k(t)$ is a rational function field, and $\mathcal{P}$ corresponds to $p(t)$, one takes $\mathcal{B} = (n-1)\cdot\infty$ for $n = \deg p(t)$, i.e. looks for a power basis $0, 1, \ldots, t^{n-1}$ of an algebraic extension $k_{\mathcal{P}} = k[t]/p(t))$ of k.

The multiplication law in the field $k_{\mathcal{P}}$ over k can be then represented in terms of the multiplication rule $L(-\mathcal{B}) \times L(-\mathcal{B}) \rightarrow L(-2\cdot\mathcal{B})$. It remains now to reconstruct functions from $L(-2\mathcal{B})$ by interpolation. For this one could take a set $\mathcal{D}$ of divisors of the first degree (places of the first degree) such that $\mathrm{card}(\mathcal{D}) > 2d(\mathcal{B})$.

This general method is summarized in the following statements from [45-46] where for an arbitrary integral (positive) divisor $\mathcal{A}$ on K, we put $k_{\mathcal{A}} = K/\mathcal{A} \stackrel{\mathrm{def}}{=} K \mathrm{mod} \mathcal{A}$.

Corollary 5.3 : Let K be a function field over k of genus $g \geq 0$ and let $\mathcal{A}$ be a prime divisor of degree $n \geq 1$ on K. Let $\mathcal{B}_0$ be a nonspecial integral divisor on K, i.e. $\dim_k L(-\mathcal{B}_0) = d(\mathcal{B}_0) - g + 1$, such that $\mathcal{B} = \mathcal{B}_0 + \mathcal{A}$ is a nonspecial divisor too. Let there be D prime divisors of first degree on K for $D > 2d(\mathcal{B})$. Then there exists a bilinear algorithm over k that computes the multiplication in the field extension $k_{\mathcal{A}}$ of k of degree n of multiplicative complexity at most $2d(\mathcal{B}) - g + 1 = 2n + 2d(\mathcal{B}_0) - g + 1$.

Let $\mathcal{A}$ be an arbitrary prime divisor on K of degree n, and $k_{\mathcal{A}}$ be its residue class field which is an extension of k of degree n. Let us denote by $\#K(k)$ the number of first degree divisors on K over k. It follows from Corollary 5.3 that whenever there exists a nonspecial divisor $\mathcal{B}_0$ on K of degree m such that $m + n \geq 2g - 1$ and $\#K(k) \geq 2m + 2n$, the multiplicative complexity, $\delta_k(k_{\mathcal{A}})$, of computation of multiplication in $k_{\mathcal{A}}$ over k, does not exceed $2m + 2n - g + 1$.

6. Diophantine Geometry with Applications to Low Complexity Algorithms.

As results of §5 show, to construct low complexity algorithm of multiplication in k-algebras, associated with function rings, one needs a supply of points (prime divisors of the first degree). Thus the main problem becomes the existence of a large number of (appropriate) rational points on curves of high genus- a classical diophantine problem.

First of all we need the existence of some auxiliary divisors for Corollary 5.3. For this we use the properties of the ζ-function $Z(u) = \zeta(K;s)$, $u = q^s$. Let us denote by $N\text{prime}_n$ the number of prime divisors on K/k of degree n. It follows that $d/du \log Z(u) = \sum_{n=1}^{\infty} u^{n-1}\{\sum_{d|n} N\text{prime}_d \cdot d\}$. We obtain:

$$q^n + 1 - \sum_{i=1}^{2g} \omega_i^n = \sum_{d|n} N\text{prime}_d \cdot d$$

for any $n \geq 1$. Since $\mid \omega_i \mid = \sqrt{q}$ for all $i = 1, \ldots, 2g$, we deduce:

a) for every $n \geq c_1 \cdot \log(g)/\log q$ there is a prime divisor of degree n on K;

b) for every $m \geq g + c_2 \log(g \cdot q)$ there exists a nonspecial positive divisor of degree m on K.

In order to have low multiplicative complexity algorithms arising from algebraic curves one sees that the remaining problem is the construction of algebraic curves over a fixed field $k = \mathbf{F}_q$ having the maximal number of divisors of the first degree for a given genus g, as $g \to \infty$.

The most general bound on the number $\#K(k)$ of points on K over k (prime divisors of first degree) is given by Weil's bound: $\mid N_k(K) - q - 1 \mid \leq 2g\sqrt{q}$. As $g \to \infty$, the upper bound in the Weil theorem is unattainable. In fact, relatively simple considerations (cf. [47]) show that

$$\lim_{g \to \infty} \frac{N_k(K)}{g} \leq \sqrt{g} - 1. \tag{6.1}$$

On the other hand, there are positive results [47-48]:

*Proposition*6.1 : Whenever q is square, the equality in (6.1) is attainable.

For example, on every classical modular curve of level m and genus g, $(m,p) = 1$, there are at least $(p-1)(g-1)$ points over $\mathbf{F}_{p^2}$.

Choosing curves over finite fields with many points on them (Fermat curves, congruence curves) we obtain a variety of new low multiplicative complexity algorithms of polynomial multiplication.

*Theorem*6.2 : Let q be a square ≥ 25. Then the multiplicative complexity $\delta_{\mathbf{F}_q}(\mathbf{F}_{q^n})$ of multiplication in the field $\mathbf{F}_{q^n}$ over $\mathbf{F}_q$ can be bounded as follows:

$$\delta_{\mathbf{F}_q}(\mathbf{F}_{q^n}) \leq n \cdot 2 \cdot (1 + \frac{1}{\sqrt{q} - 3}) + o(n)$$

as $n \to \infty$.

For $q = 2$ we have the following upper and lower linear bounds on multiplicative complexities (see §4 and the multiplication rule (5.10)):

$$3.52 \cdot n \leq \delta_{\mathbf{F}_2}(\mathbf{F}_{2^n}) \leq 6n \tag{6.2}$$

for a large n.

For $\mathbf{F}_2$ and moderate n and m (below 100) the best algorithm of polynomial multiplication over $\mathbf{F}_2$ has multiplicative complexity of linear character with a constant around 4.

All algorithms of Theorem 6.2 and (6.2) are constructive (one can construct appropriate algebraic curve codes in polynomial time).

The relationship with the linear coding theory, indicated in §5, extends further. According to §5, with every multiplication algorithm one can associate an appropriate code, and low minimal complexity algorithms give rise to very good codes. Our algorithms (Theorem 6.2 and others [45-46]) lead to Goppa-like linear codes arising from algebraic curves over finite fields [48].

New algorithms for polynomial multiplication (convolution) over finite fields are to be extended to algorithms over $\mathbf{Z}[1/2]$ or $\mathbf{Z}$ to be particularly useful in applications to bignum computations. Here we encounter a variety of diophantine problems connected with the number of solutions of classical diophantine equations, see [49-50].

Apparently, one of the conditions that guarantees the existence of fast polynomial multiplications over $\mathbf{Z}$, is the existence of algebraic number fields K, $[K : Q] = n$ with a (large) number M of "exceptional units". We say that $\epsilon_1, \ldots, \epsilon_M$ are "exceptional units" in K if $\epsilon_1 = 1$ and $\epsilon_i - \epsilon_j$ is a unit in K for $i \neq j$. This concept first appears in works of Minkowski and Hurwitz. Recently it was studied by Lenstra. In particular, Lenstra had shown that the following explicit bound

$$M > \frac{n!}{n^n}(\frac{4}{\pi})^s \mid D_K \mid^{\frac{1}{2}},$$

(where D_K is the discriminate of K and s is the number of complex archimedean primes) on the number M of exceptional units implies the Euclidianity of K with respect to the norm of K.

Obvious bound on M for an arbitrary field K of degree N is:

$$2 \leq M \leq 2_n.$$

On the other hand, by looking at subfields of cyclotomic fields and subfields of Abelian extensions of quadratic imaginary fields, one can construct infinite families of fields K of degree n such that

$$M = O(n(log \log n)^2)$$

for the number M of exceptional units in K.

Another similar object, also connected with Euclidianity, is useful in the construction of fast polynomial multiplication algorithms over ring of constants $\mathbf{Z}[1/2]$. In this case,

one considers a sequence of elements $\delta_i : i = 1, \cdots, M$ in K such that $\delta_i - \delta_j$ are (nonzero) divisors of powers of 2. Such sequences and corresponding polynomials were introduced by H. Cohn [51] as "dyadotropic".

Another interesting arithmetic problem arising from the construction of fast algorithms is the existence of long sequences of "relatively prime over $\mathbf{Z}$" polynomials. In this problem, one is looking at a sequence $(Q_\alpha(x) : \alpha \in A)$ of polynomials from $\mathbf{Z}[x]$ (with, say, leading coefficient 1) such that for $\alpha \neq \beta$,

$$res(Q_\alpha, Q_\beta) = 1$$

(i.e $BQ_\alpha + CQ_\beta \equiv 1$ for B, $C \in \mathbf{Z}[x]$). One can understand these sequences better by looking at reductions mod$p$ (particularly for $p = 2$) and comparing lists of irreducible polynomials over $\mathbf{F}_p[x]$. Ideal sequences should be long with low degrees of Q_α. As with Euclidian fields, the best source of construction of such sequences is by means of division polynomials, particularly for elliptic curves with complex multiplication. The corresponding elliptic divisibility polynomials for an elliptic curve $\mathbf{E}/\mathbf{Q}$ with complex multiplication in an imaginary quadratic field L (with the class number, say, one) are given by:

$$\psi_\mu(u) = \frac{\theta(\mu u)}{\theta(u)^{Norm(\mu)}}$$

where $\mu \in O$ (O is an order in L), and $\theta(u)$ is a normalized σ – function (corresponding to the lattice $\Lambda = \Omega O$ in $\mathbf{C}$:

$$\theta(u) = \sigma(u) e^{-s_2(\Lambda)u^2/2},$$

$$s_2(\Lambda) = \lim_{\epsilon \to 0^+} \sum_{\omega \in \Lambda, \omega \neq 0} \omega^{-2} |\omega|^{-2\epsilon}.$$

For moderate sizes of A, we had constructed sequences of relatively prime polynomials over $\mathbf{Z}$ explicitly.

Our new algorithms of polynomial multiplication over $\mathbf{Z}$ for small n ($n \leq 100$) are matching in minimal complexity algorithms over $\mathbf{F}_2$. For arbitrary n, we can improve known upper bound of the minimal complexity, $\mu_{\mathbf{Z}}(n, n)$ of multiplication of polynomials of degree $n-1$ with the ring $\mathbf{Z}$ of constants by a factor $\log^2 \log \log n$.

Development of entirely new classes of minimal complexity algorithms is closely connected with the study of arithmetic surfaces (a la Arakelov-Faltings). The development of our interpolation methods on algebraic manifolds strongly suggests a conjecture, that all polynomial multiplication algorithms are reducible to combinations of valuation algorithms. These valuation algorithms correspond to archimedean and nonarchimedean valuations on appropriate algebraic surfaces. Classical algorithms (including high-school method, ordinary interpolation algorithms, residue number systems, FFT and number-theoretic FFT) all correspond to projective spaces $k\mathbf{P}^n$. An open problem on whether $\mu_{\mathbf{Z}}(n, n) = O(n)$ is equivalent in this context to special properties of L–functions of Abelian varieties at $s = 1$.

Let us look now at the total algebraic complexity of polynomial multiplication algorithms. First of all, there are yet no satisfactory definitions of total (multiplicative,

additive, scalar, etc.) complexity, corresponding to the true hardware organization. Interesting definitions and results belong to V. Pan. In line with his definitions, we can give some lower end bounds on the total algebraic complexities of polynomial multiplication. For this we consider an arbitrary (commutative) scheme having (i) μ nonscalar (essential) multiplications; (ii) α chain additions, and (iii) σ scalar multiplications. We have to assign a weight to each scalar multiplication. If c is a scalar which is an algebraic number, its size $M(c)$ (or height) is usually determined as

$$M(c) = \prod_v max(1, \|c\|_v),$$

where v runs through all valuations of the smallest field K containing c. E.g. for a rational $c = p/q \neq 0$ with $(p, q) = 1$, $M(c) = max(|p|, |q|)$.

It is not unreasonable to attach to the scalar multiplication by c a weight $\log M(c)$ (this is an understatement of a hardware performance).

The total "cost" of the computational scheme can be represented as

$$C = \mu + \alpha + \sum_c \log M(c),$$

were μ and α are the numbers of multiplication and addition chains, and c runs through all scalar multiplication chains.

A lower bound on C defined this way in the problem of multiplication of two polynomials of degree $n - 1$ is *non-linear* in n:

$$C \geq n \log_2 n.$$

On the other hand, in our new low multiplicative complexity algorithm presented here, the "total cost" C is always

$$C = O(n \log n),$$

which makes these algorithms as attractive in applications as any of Fast Fourier methods.

To decrease the additive complexity (and cost of scalar multiplications) one has to examine the choice of evaluations or interpolation points (choices of valuations, archimedian and nonarchimedian). These choices are determined by the hardware realization, were the binary structure of numbers plays a crucial role, e.g. multiplications by powers of 2 amount to simple shifts. This problem appeared initially in polynomial interpolation and is described in [52], see also [53]. Explicitly we can ask of the cost of evaluation of n-th degree polynomial at $n + 1$ points, $\{x_i\}$ (over an appropriate ring of scalars, typically a field). Following [52], evaluation of a set of points x_i is called "fast", if it can be accomplished in $O(n \log n)$ operations. The FFT algorithms provide with sets of points, where fast evaluation can be accomplished. These points are roots of unity of a given order, which is highly composite, typically a power of 2. In any ring, were such roots of unity exist (and the order of roots can be inverted), fast evaluation is possible. It is this fast evaluation by means of FFT that gives rise to fast convolution and then fast interpolation algorithms based on fast convolutions.

We developed recently a variety of new fast algorithms of rational function interpolation and evaluation the are not constrained by restrictions of (number-theoretic) FFT relating the word size and the length of the transformation.

To describe our algorithms we look now at a general rational function interpolation problem. In the partial fraction representation of rational functions one looks at the rational function with poles only at given (distinct) points: $\alpha_1, ..., \alpha_n$. The general form of such a rational function is

$$R(z) = \sum_{i=1}^{n} \frac{x_i}{z - \alpha_i}. \tag{6.3}$$

Thus its divisor is $(-\sum_{i=1}^{n} \alpha_i + \infty)$. The evaluation problem for this function (or its divisor), consists of simultaneous determination of n values of $R(z)$ at $z = \beta_1,...,\beta_n$ (distinct from α_i):

$$y_j = R(z)|_{z=\beta_j} = \sum_{i=1}^{n} \frac{x_i}{\beta_j - \alpha_i} : j = 1,..n. \tag{6.4}$$

One can consider the transformation from x_i in (6.3) to y_j in (6.4) as a discrete (rational) transform determined by two divisors or $\mathcal{A} = (\sum_{\backslash=\infty}^{\backslash} \alpha_\backslash)$ and $\mathcal{B} = (\sum_{|=\infty}^{\backslash} \beta_|)$. It is easy to see that the inverse to this transformation is the following explicit one:

$$x_i = -\sum_{j=1}^{n} \{\frac{P_{\mathcal{B}}(\alpha_i)}{P'_{\mathcal{A}}(\alpha_i)} \frac{P_{\mathcal{A}}(\beta_j)}{P'_{\mathcal{B}}(\beta_j)}\} \frac{y_j}{\beta_j - \alpha_i} : \tag{6.5}$$

$i = 1, ..., n$. Here, $P_{\mathcal{A}}(z)$, $P_{\mathcal{B}}(z)$ are polynomials of degree n having as roots $\{\alpha_i\}$, and $\{\beta_i\}$, respectively.

Into the scheme (6.3-5) fall discretizations of important one- and multi-dimensional integral transform (with a variety of quadrature approximations methods). Among singular integral transformations that can be described by direct and inverse schemes (6.4-5) the most obvious is the Hilbert transform on a circle, its proper discretization, and computations via FFTs are described in detail in [55]. A finite Hilbert transform corresponds in the scheme (6.4-5) to α_i being N-th roots of 1, and β_j being N-th roots of -1.

We present now several classes of fast rational evaluation algorithms with computational cost $O(n \log n)$, that generalize finite Hilbert transforms. These new transformations correspond to a variety of singular integrals taken over one-dimensional complex continuum and to singular integrals over fractal sets, and to singular integrals with elliptic function kernels. The latter transformations have interesting number theoretic and modular interpretation and we refer to them as Fast Elliptic Number Theoretic Transforms (FENTT).

To describe all transformations of this class that can be evaluated in $O(n \log n)$ operations, we look at α_i and β_j given as roots of polynomials $P_{\mathcal{A}}(z)$ and $P_{\mathcal{B}}(z)$, respectively, where polynomials $P_{\mathcal{A}}$ and $P_{\mathcal{B}}$ correspond to iteration of (fixed) polynomials and rational functions. Thus we start with a sequence of degrees (radices) $D_1, ..., D_m$ and with rational

functions (polynomials) $R_1(z),...,R_m(z)$ of degrees $D_1,...,D_m$, respectively. The polynomials $P_\mathcal{A}$ and $P_\mathcal{B}$ have degrees $n = D_1...D_m$ and have roots as preimages of two distinct points α and β under iterated mappings

$$z_{i+1} = R_i(z_i).$$

Thus $P_\mathcal{A}(z)$ is defined as (the numerator of) $R_m(R_{m-1}(...(R_1(z))...)) = \alpha$, and $P_\mathcal{B}(z)$ as (the numerator of) $R_m(R_{m-1}(...(R_1(z))...)) = \beta$.

The most interesting case is that of $D_1 = ... = D_m = D$ and $R_1 = ... = R_m = R$ being a fixed rational function of degree D. The fast algorithm to compute (6.4-5) is similar in the flow diagram to the mixed radix FFT schemes.

The mixed radix FFT algorithm is a special case with $R_i(z) = z^{D_i} : i = 1,...m$. The total computational cost of such mixed degree algorithm depends on the computational cost of multiplication of polynomials of degrees $O(D_i)$. In case of $D_i = O(D)$, and of large m, the total computational cost of evaluation of (6.4-5) is $O(mD_1...D_m)$.

Among rational transforms that fit into this scheme we can mention: a) Hilbert-like transforms for various Julia sets corresponding to polynomial or rational mappings $z \to R(z)$; b) singular integral transformations with elliptic function kernels. In the case a), contours of integration can have arbitrary fractional dimension. In the case b), a continuous analog of the transformation (6.5) is the following

$$Y(s) = \int_0^\omega X(t)\wp(t-s)dt$$

for the Weierstrass elliptic function $\wp(u)$. The latter transform and its discrete versions are particularly well-suited for arithmetic interpretations. If an elliptic curve **E** is defined over a finite field $k = \mathbf{F}_p$ (e.g. is a reduction mod p of an elliptic curve over **Q**), then the set of its k-rational points is an Abelian group of order $N_p = p - a_p + 1$ for $|a_p| \le 2\sqrt{p}$. Moreover, for any integral a, $|a| < 2\sqrt{p}$, there is an elliptic curve over $\mathbf{F}_p$ with $N_p = p-a+1$. Whenever 2^n divides N_p, one has points of order 2^n on an elliptic curve, all defined over $\mathbf{F}_p$. Consequently, the fast evaluation algorithm can be applied in this case with $D_1 = ... = D_m = 4$. The rational function $R(z)$ in this case is the duplication formula for x-coordinate in the Weierstrass cubic form of an elliptic curve E: $y^2 = 4x^3 - g_2x - g_3$:

$$R(x) = -2x + \frac{(6x^2 - g_2/2)^2}{4y^2}$$

The choice of α and β should be of x-coordinates of second order points on **E**. Whenever prime p is such that p is within distance $2\sqrt{p}$ from a power of 2 (or from a highly composite number) one has a very fast algorithm of rational evaluation and transformation of length $O(p)\mathrm{mod}p$.

The same method can be used for a composite number of M if one chooses an appropriate elliptic curve over $\mathbf{Z}/M\mathbf{Z}$, whose reduction mod p for prime factors p of M have highly composite Abelian group of $\mathbf{F}_p$-rational points. Using the standard facts of the distribution of highly composite numbers, see [56], we conclude that with any number M we have FENTT of length $O(M)$ over $\mathbf{Z}/M\mathbf{Z}$ with computational cost $O(M \log M)$. FENTT algorithms are particularly attractive in parallel implementation.

7. Evaluation of Solutions of Linear Differential Equations.

Recently we have developed new fast algorithms of power series computations for important classes of algebraic functions and solutions of linear differential equations with rational function or algebraic coefficients, [32], [36], [27]. These algorithms provide also with efficient methods of evaluation and of analytic continuation of solutions of these equations. Our algorithms are all based on the analysis of (linear) recurrences on coefficients of power series expansions.

We differentiate between the operational and the total (boolean) complexity. By operational complexity of an algorithm one understands the number of primitive operations (most notably additions and multiplications), independent of the sizes of numbers involved, needed to complete this algorithm. By the total (or boolean) complexity we understand the *total* number of primitive operations (on short or single-bit data) needed to complete a given program. The main distinction between the conversion from the operational to the total complexity, depending on the size of numbers involved, is described by the total complexity of multiplication of big numbers.

Let us denote by $M(n)$ the total complexity of multiplication of two n-bit integers. Then the best known upper bound on $M(n)$ belongs to Schonhage-Strassen [57]:

$$M(n) = O(n \log n \log\log n).$$

In comparison, a total complexity of addition is relatively simple: it is only $O(n)$.

All algebraic operations on bigfloats have boolean complexity of the same order of magnitude as a multiplication. For example, if $B(n)$ denotes one of the following total complexities: division of n-bit bigfloat numbers, square root extraction, or raising to the fixed (rational) power, then $B(n) = O(M(n))$, and $M(n) = O(B(n))$.

Our methods based on fast computation of solutions of matrix linear recurrences allow us to construct new algorithms of low total complexity for evaluation of solutions of linear differential equations with an arbitrary precision. Such algorithms of low total complexity were till now unknown for general special functions. Low complexity algorithms were known only for algebraic function computations (standard Newton method), and for elementary functions.

We describe now our new algorithms [27, 36] of low total complexity for computation of values of solutions of arbitrary differential equations, neither related to any rapidly convergent analytic transformations, nor to any low operational complexity methods. These methods differ in boolean complexity at worst by factor $\log^2 n$ or $\log^3 n$ from the boolean complexity $M(n)$ of algebraic computations.

Let us look at an arbitrary linear differential equation with rational (polynomial) coefficients, either in the scalar form

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0, \qquad (7.1)$$

or, in the general matrix form,

$$\frac{d}{dx}Y(x) = A(x) \cdot Y(x), \tag{7.2}$$

where $a_i \in \mathbf{C}(\mathbf{x})$, and $A(x) \in M_n(\mathbf{C}(\mathbf{x}))$. We are interested in the evaluation of solutions of (7.1) and (7.2) with arbitrary precision using the method of (formal) power series expansions of [36]. Having specified initial conditions for $y(x)$ or $Y(x)$ at $x = x_0$, we want to evaluate within a given precision l $y(x)$ or $Y(x)$ at another point $x = x_1$. For all practical purposes we assume that values of x_0 and x_1 are given correctly with the precision of l bits, or as rational or algebraic numbers of sizes less than l. To determine values at $x = x_1$ from those at $x = x_0$ one has to specify a path γ from x_0 to x_1 on the Riemann surface (or its universal covering) of $y(x)$ or $Y(x)$, see [27].

The global recipes [27, 36] for the analytic continuation of $Y(x)$ along γ are combined with better local methods of evaluation of $Y(x)$ from $x = x_0$ to $x = x_1$. For this we are using local power series of $Y(x)$ to continue $Y(x)$ from $x = x_0$ to $x = x_1$ making several steps between x_0 and x_1, releasing consecutive blocks of x_1 in bursts. We call this method "bit-burst" method. This way we arrive at the following general theorem that gives an upper bound on the total (boolean) complexity:

Theorem 7.1. Let (7.2) be a given linear differential equation with rational function coefficients, and $Y(x)$ be its arbitrary (regular) solution with initial conditions at $x = x_{in}$, where x_{in} is an K-bit number. Given a path γ from x_{in} to an K-bit number x_{fin} (on the Riemann surface of $Y(x)$) of length L, one can evaluate $Y(x)|_{x=x_{fin}}$ at $x = x_{fin}$ with the full K-bit precision at most

$$O(M(K)(\log^3 K + \log L))$$

bit operations.

The bit-burst method in the general form of Theorem 7.1 is not the best possible: one would like to see $\log^3 K$ replaced by $\log K$ always. Sometimes the complexity can be lowered:

I. If x_{in} and x_{fin} are fixed rational numbers, then the computation of $Y(x)$ with the full K bits of precision has total (bit) complexity at most $O(M(K)(\log^2 K + \log L))$.

II. If the differential equation (7.1) or (7.2) possesses special arithmetic properties, bit-complexity can be lowered. E.g. if the (7.1)-(7.2) possesses a solution which is either an E-function or a G-function, then the general bit bound of Theorem 7.1 can be lowered to

$$O(M(K) \cdot (\log^2 K + \log L)).$$

If, further, like in x_{in} and x_{fin} are fixed rational numbers, and $Y(x)$ is built from E- or G-functions, then K significant digits of $Y(x_0)$ can be computed in

$$O(M(K) \cdot (\log K + \log L))$$

bit operations.

This last bound is unsurpassed by any other algorithms even for elementary functions, like the exponent, where low operational complexity algorithms are well known, see [58].

Our original interest in the development of power series evaluation facilities was purely transcendental. We wanted to have the ability to compute the monodromy group of linear differential equations with an arbitrary precision to check various hypotheses of transcendence, algebraicity and the existence of algebraic relations among elements of the monodromy matrices. The "constants" appearing as elements of monodromy matrices encompass classical constants of geometry and analysis. Among them there are periods of algebraic varieties including values of the Euler Γ- and B-functions (at rational points) and other integrals of elementary and algebraic functions over closed paths.

To compute the monodromy group of linear differential equations we use the power series method based on the direct analytic continuation of a fundamental system of solutions. In this method one starts with a fundamental system of solutions

$$\overline{Y} = (y_1(x), \ldots, y_n(x))$$

of (7.1) [given by their initial conditions] and analytically continues it along a closed path with no singularities on it. The system Y analytically continued along a path γ undergoes a linear transformation

$$Y \mapsto_\gamma Y \cdot M(\gamma).$$

The set of all matrices $M(\gamma)$ is a monodromy group of (7.1).

Areas of applications include:

1) multiple precision computation of Abelian integrals and their periods. These computations are used then in the transcendental solution of the problem of reduction of Abelian integrals. In order to determine the reducibility of Abelian integrals to the lower genra (e.g. when an Abelian integral is an elementary function) one looks at the **Z**-relations between periods.

2) Another application of monodromy computations is to the solution of the direct and inverse Galois problem, when one wants to find a Galois group of a given algebraic function field (differential equation) or wants to construct a field with a given Galois group. Our package is designed mainly for the direct Galois problem, but we found it extremely convenient to use for solution of the inverse Galois problem when the number of parameters is not large and one can numerically invert the function (monodromy matrix output) generated by our program.

3) Uniformization theory seems to be an attractive proving ground for application of monodromy packages. For us the crucial problem was the question of arithmetic nature of parameters in the solution to the uniformization problem, including: a) algebraic equations (i.e. their coefficients) defining Riemann surfaces to be uniformized; b) invariants of discrete groups that uniformize these Riemann surfaces (Fricke parameters and more sophisticated parametrizations of Teichmüller spaces); and finally, the most notorious group of parameters c) assessory parameter that uniquely determine the differential equation of the second order, ratio of solutions of which determines the inverse to the uniformizing function.

Interesting applications of our numerical studies of uniformization theory include massive computations of accessory parameters (and corresponding Fuchsian groups) for hyperelliptic surfaces. One of the results of the computations [32] was the negative solution to the Whittaker conjecture (1928), which predicted the explicit expression of accessory parameters for hyperelliptic surfaces of genus $g \geq 2$, by looking at their birational transformations. The true structure of Teichmuller spaces, as revealed by our numerical computations is quite complex [27], [32].

4) Another group of applications of explicit solutions of accessory parameter problem involves explicit determination of conformal mappings of complicated domains using appropriate linear differential equations. This gives a high-precision method that can be used in adaptive grid computations.

Complexities of power series computations can be also investigated from the point of view of parallel (vector) implementation. The parallel (vector) methods are important because they seem to be the only way to address large jobs. Algorithms that compute power series coefficients and values of functions by means of linear recurrences are particularly well-suited for various vector and parallel implementations.

The methods that we propose, bit-burst algorithms for computations of arbitrary linear differential equations, have always depth $O(\log n \log\log n)$ even though, in general, the total bit operation count is $O(M(n) \cdot \log^3 n)$. In fact, for E- and G-functions the total bit operation count with the same depth is only $O(M(n) \cdot \log^2 n)$.

8. Polynomial Root Finding

One calls a (univariant) polynomial sparse, if it has a number of nonzero coefficients significantly less than its degree.

Sparse polynomials, particularly trinomials and quadrinomials, are extremely interesting from many points of view. Let us mention now just one problem: which Galois groups G can be realized as Galois groups of trinomials $ax^n + bx^m + c = 0$ (for $a, b, c \in \mathbf{Z}$ or for $a, b, c \in \mathbf{Z}[\mathbf{x}]$)?

The generic Galois group of a trinomial of degree n is S_n, so roots of a trinomial cannot be determined in closed form for $n \geq 5$. Nevertheless one can express the roots of trinomials in a closed form as infinite series with binomial coefficients. Such an expression can be found in Ramanujan's writing.

Ramanujan normalizes the trinomial equation in the following form

$$aqx^p + x^q = 1,\ a > 0 \text{ and } 0 < q < p.$$

For $n > 0$ and a particular root of a trinomial $aqx^p + x^q = 1$, Ramanujan [R1,Ch.3 of his Second Notebook] proves the expansion:

$$x^n = \frac{n}{q} \cdot \sum_{N=0}^{\infty} (-qa)^N \cdot \frac{\Gamma(\frac{n+pN}{q})}{N!\Gamma(\frac{n+pN}{q} - N + 1)}. \tag{8.1}$$

This expansion converges when $|a| \leq p^{-p/q} \cdot (p-q)^{(p-q)/q}$.

This expansion of Ramanujan is usually attributed to Lambert (1758). The usual derivation of this formula is based on Lagrange inversion theorem (1770). References to this and other similar Ramanujan's formulas can be found in Brendt, Evans, Wilson, [59].

"Ramanujan's" formula can be also derived directly from the trinomial equation using the transformation of contour integrals into Barn's type integrals of Γ-functions. This derivation had been achieved by Hj. Mellin [60], and generalized by him to arbitrary algebraic equations.

These integral representations are particularly convenient for sparse polynomials, where they also can be converted into multivariate power series expansions.

Let us look at an arbitrary $k+2$-nomial

$$x^n + a_1 x^{n_1} + a_2 x^{n_2} + \ldots + a_k x^{n-k} - 1 = 0, \tag{8.2}$$

$n > n_1 > n_2 > \ldots > n_k > 0$. Then we have the following expansion of all n roots x_j of this equation, or of their $\mu - th$ powers x_j^μ:

$$\epsilon_j^\mu \cdot x_j^\mu = 1 + \mu \sum_{N=1}^{\infty} \frac{(-1)^N}{n^N} \cdot \sum_{j_1+\ldots+j_k=N} A_{j_1,\ldots,j_k} \cdot a_1^{j_1} \ldots a_k^{j_k}, \tag{8.3}$$

where one puts:

$$A_{j_1,\ldots,j_k} = \frac{\prod_{i=1}^{N-1}(\mu - n_i + j_1 n_1 + \ldots + j_k n_k)}{j_1! \ldots j_k!}.$$

The series (8.3) converge always whenever

$$max(|a_i| : i = 1, \ldots, k) < min(n/(k\sqrt{[n]}n_1^{n_1}(n-n_1)^{n-n_1}), n/(k\sqrt{[n]}n_k^{n_k}(n-n_k)^{n-n_k})).$$

In principle, the power series formulas like Ramanujan (8.1) or Mellin's (8.3) can be used directly for the root finding of sparse polynomials. It is much more practical, though to use our methods of analytic continuation of power series solutions of linear differential equations satisfied by algebraic functions. These differential equations can be represented in terms of equations on multidimensional generalized hypergeometric functions. These equations can be integrated and analytically continued everywhere. The most general expression for a system of Fuchsian linear differential equations on x_i^μ, as functions of $a_1, \ldots, a_k$, can be deduced from the power series representation of (8.3) for x_j^μ.

Namely, the (algebraic) functions x_j^μ for roots x_j of (8.2) satisfy the following system of k Fuchsian linear differential equations on $\partial/\partial a_1, \ldots, \partial/\partial a_k$:

$$\{(-1)^{n_i} \cdot n^n \cdot \frac{\partial^n}{\partial a_i^n} - \sqcap_{m=0}^{n_i-1}(n_1 a_1 \frac{\partial}{\partial a_1} + \ldots + n_k a_k \frac{\partial}{\partial a_k} + \mu + nm)$$

$$\times \sqcap_{m=0}^{n_i'-1}(n_1' a_1 \frac{\partial}{\partial a_1} + \ldots + n_k' a_k \frac{\partial}{\partial a_k} - \mu + nm)\} x_j^\mu = 0; n_i' = n - n_i : i = 1, \ldots, k. \quad (8.4)$$

This system (8.4) of linear differential equations can be integrated starting from $a_1 = a_2 = \ldots = a_k = 0$ with initial conditions $x_j = \epsilon_j$ for $n-th$ roots of unity $\epsilon_j^n = 1 : j = 1, \ldots, n$. Alternatively integration can start at ∞, or at any other point in k- space, where x_j are known.

The methods of analytic continuation of algebraic functions x_j from $a_i = 0$ can be easily implemented, and its only nontrivial part concern the regular (Puiseux) power series expansions of x_j in the neighborhood of singularities of differential resolvent equations (8.4), see [61]. This leads to an iterative algorithm, where to compute the roots of $k+2$ -nomial one has to precompute the roots of $k+1$-nomials (that give the singularities of the former branches of algebraic functions), etc. This method is perfect for parallel implementation, since on each level of iteration all roots are computed independently. The storage requirements are determined only by the number of roots one wants to see. Moreover, this iterative method was implemented in vector and array hardware, e.g. on IBM 3090-VF or on CRAYs, because most of the operations are vectorizable loops.

The complexity bound for the computation of roots of sparse polynomials using the analytic continuation of solutions of differential resolvents is the following.

Theorem 8.1. One can compute all n roots of a k-nomial of degree n with the precision M (of leading digits) in $O(k^2 log^2 M)$ parallel steps on n processors. If $k \ll n$, then one can compute $m(\leq n)$ roots of a k-nomial of degree n with the precision of M (digits) in $O(k\, log^2 M)$ parallel steps on $O(m)$ processors.

The crucial osbtacle in the complex root finding for large degrees of polynomials is the need for multiple precision computations. Unfortunately high precision requirements make the programming of the root finding methods awkward in any vector or parallel environment. Because of these constraints our fast polynomial root finding algorithms

for *dense* polynomials involved degrees only in thousands. The largest "random" dense polynomial we completely solved in double precision on a single processor of CRAY II had degree 15,000.

The target of our large degree polynomial root finding programs was the analysis of the distribution of complex roots of a real (or complex) polynomial with random coefficients. Specifically we looked at normally distributed random coefficients, though various other distributions were analyzed as well. The distribution of real roots of a random polynomial of degree n (there are $O(logn)$ of them) was described by Hardy-Littlewood-Kac. Much less is known about the distribution of the complex roots in addition to the obvious statement that "most of the roots are uniformly distributed around the unit circle."

We have conducted extensive computations of complex roots for large series of random polynomials (normal, uniform, uniform $\{0,1\}$, and other distribution of coefficients) of degrees varying from 500 to 15,000. While degrees up to 1,000 are easy to handle on a PCAT with an accelerator board (in our case it was DSI-020), degrees higher than 1,000 required a vector facility. We used the vector facility of IBM 3090-VF for degrees up to 7,000, and CRAY II, with time provided by the NSF Grant, for degrees 7,000-10,000 with the maximum of 15,000. The limit on degrees is a consequence of precision constraints, and an increase in precision significantly inhibits the vectorization and slows down the computations. We present some of the outputs of the pictures of complex roots of random polynomials with normal distribution of coefficients, with parts of the picture blown up for detail.

These computations were based on our new parallel algorithms of root finding for dense polynomials.

The crucial problem in the construction of the truly polynomial time root finding methods is the ability to deal with possible clustering of roots.

We propose a new (probabilistic) method based on the analytic continuation of algebraic functions through their singular points. This method uses in an indirect way the homotopy method, the Euler-Newton method and the Lobachevsky method. Our method is based on the deformation of a polynomial $P(x) \in \mathbf{Z}[\mathbf{x}]$ $degP = n$, into a bundle $P(x,t)$ with the initial position $P(x,0) = P_0(x)$ for a random polynomial $P_0(x)$, and with the final position $P(x,1) = P(x)$. We then study the algebraic functions $x_i = x_i(t)$ that are branches of an algebraic function defined by the equation

$$P(x,t) = 0.$$

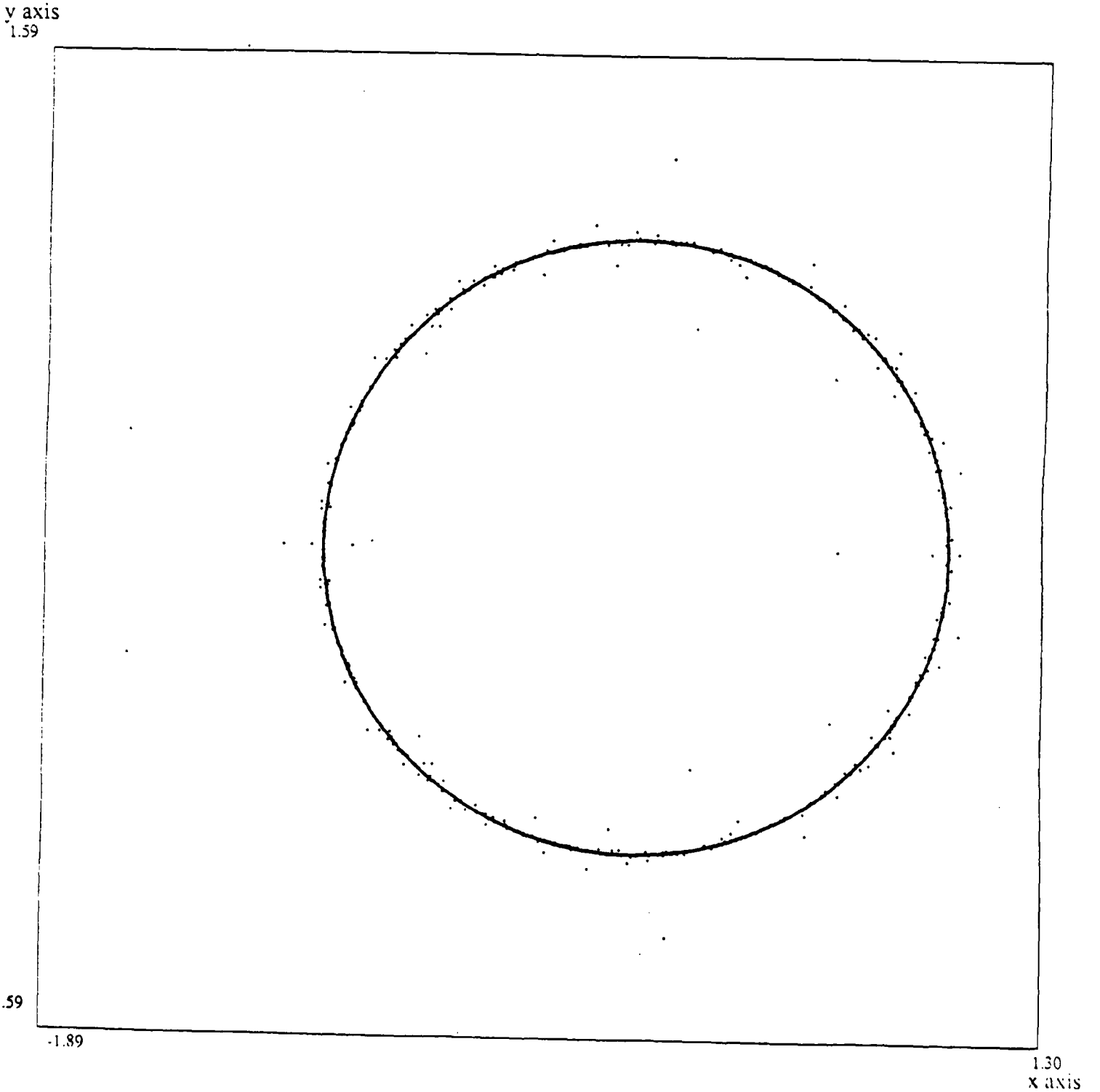
Theorem 8.2. If $P(x) \in \mathbf{Z}[\mathbf{x}]$ has degree n, then all n zeroes of $P(x)$ can be found with the precision M of leading digits ($M \gg 1$) in $O(M \log n)^{O(1)}$ parallel steps on $n^{O(1)}$ processors. For a generic polynomial $P(x)$ of degree n, all n complex zeroes of $P(x)$ can be found with the precision of M leading digits ($M \gg 1$) in $O(\log M \log n)^{O(1)}$ parallel steps on $n^{O(1)}$ processors.

Plot Shows Complex Zeros of the Polynomial

The Degree of the Polynomial Is 7000

This is a random polynomial. Seed is -1

This is a Whole Root Picture

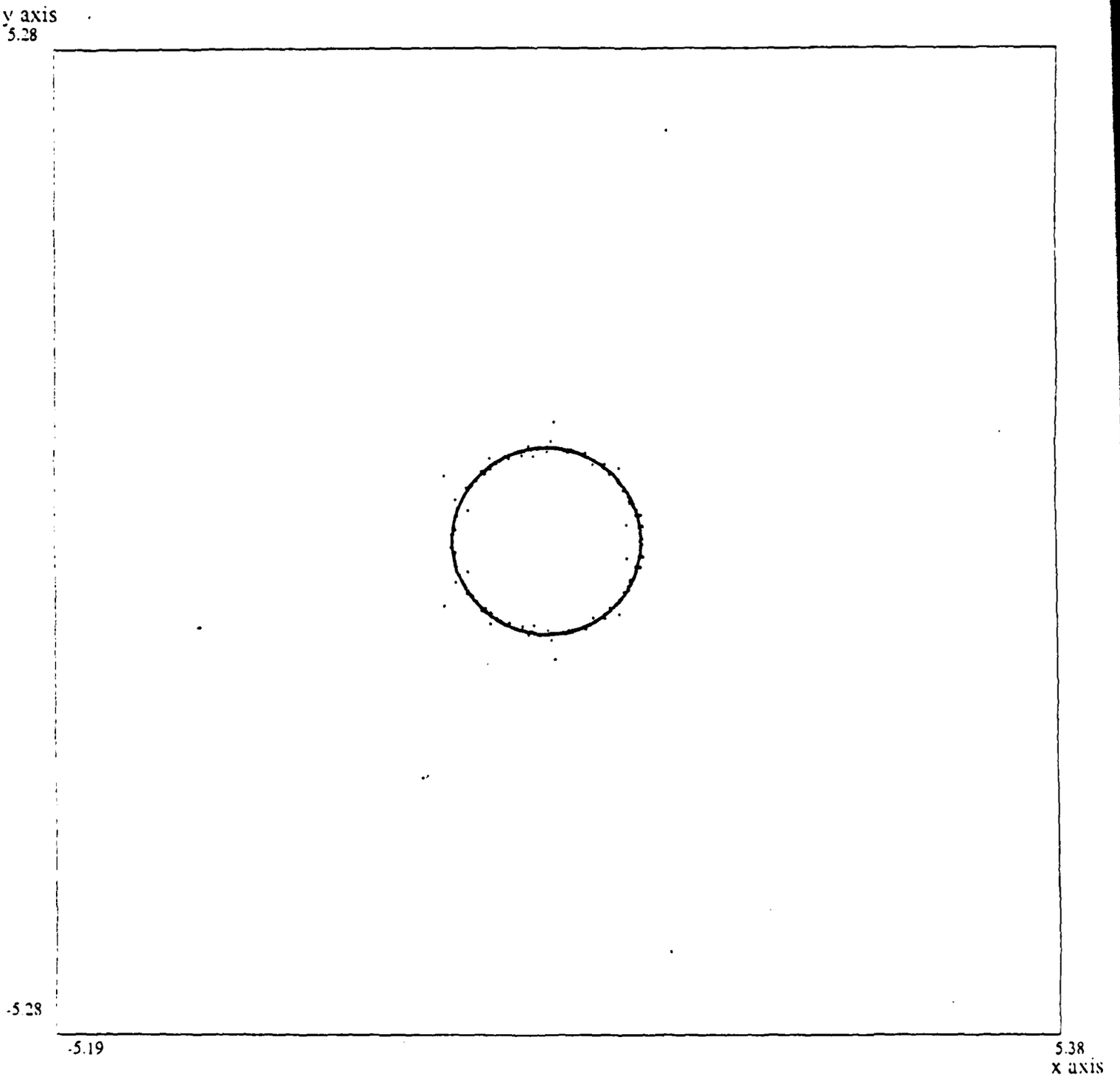


Plot Shows Complex Zeros of the Polynomial

The Degree of the Polynomial Is 7999

This is a random polynomial. Seed is -3

This is a Whole Root Picture

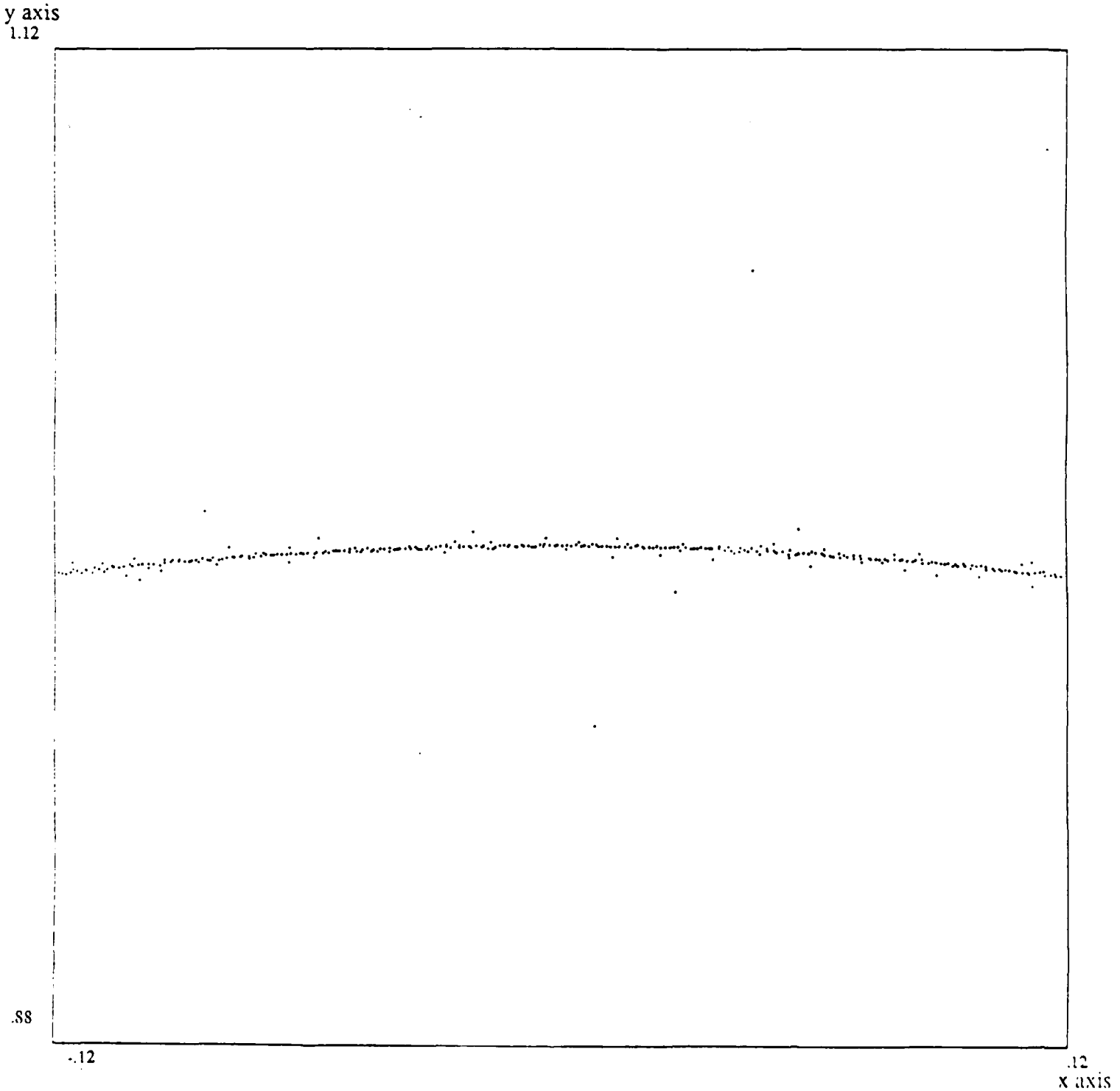


Plot Shows Complex Zeros of the Polynomial

The Degree of the Polynomial Is 7999

This is a random polynomial. Seed is -3

This is a Window of the Whole Root Picture



REFERENCES.

[1] P. Griffiths, Periods of integrals on algebraic manifolds: summary of main results and discussion of open problems, Bull. Amer. Math. Soc., 75 (1970), 228-296.

[2] C.L. Siegel, Über einige Anwendungen diophantischer Approximationen, Abh. Preuss. Akad. Wiss. Phys. Math. Kl., 1, 1929.

[3] C.L. Siegel, Transcendental Numbers, Princeton University Press, Princeton, 1949.

[4] A.B. Shidlovsky, The arithmetic properties of the values of analytic functions, Trudy Math. Inst. Steklov, 132 (1973), 169- 202.

[5] A. Baker, Transcendental Number Theory, Cambridge University Press, Cambridge, 1975.

[6] G. V. Chudnovsky, On some applications of diophantine approximations, Proc. Nat'l. Acad. Sci. USA, 81 (1984), 1926- 1930.

[7] W. M. Schmidt, Diophantine Approximations, Lecture Notes Math., v. 785, Springer, N.Y. 1980.

[8] A.L. Galočhkin, Lower bounds of polynomials in the values of a certain class of analytic functions, Mat. Sb., 95(1974), 3936-417.

[9] G. V. Chudnovsky, Padé approximations and the Riemann monodromy problem, in Bifurcation Phenomena in Mathematical Physics and Related topics, D. Reidel, Boston, 1980, 448-510.

[10] G. V Chudnovsky, Measures of irrationality, transcendence and algebraic independence. Recent progress, in Journees Arithmetiques 1980 (Ed. by J.V. Armitage), Cambridge University Press, 1982, 11-82.

[11] E. Bombieri, On G-functions, in Recent Progress in Analytic Number Theory (Ed. by H. Halberstram and C. Hooly), Academic Press, N.Y., v. 2, 1981, 1-67.

[12] K. Väänänen, On linear forms of certain class of G-functions and p-adic G-functions, Acta Arith., 36(1980), 273-295.

[13] G. V. Chudnovsky, On applications of diophantine approximations, Proc. Nat'l. Acad. Sci. USA, 81 (1984), 7261- 7265.

[14] D.V. Chudnovsky, G.V. Chudnovsky, Applications of Padé approximations to diophantine inequalities in values of G- functions, Lecture Notes Math., v. 1135, Springer, N.Y., 1985, 9- 51.

[15] T. Honda, Algebraic differential equations, Symposia Mathematica, v. 24, Academic Press, N.Y., 1981, 169-204.

[16] D. V. Chudnovsky, G.V. Chudnovsky, Applications of Padé approximations to the Grothendieck conjecture on linear differential equations, Lecture Notes Math., v. 1135, Springer, N.Y. 1985, 52-100.

[17] D.V. Chudnovsky, G.V. Chudnovsky, Padé approximations and diophantine geometry, Proc. Nat'l. Acad. Sci. USA, 82(1985), 2212-2216.

[18] J.-P. Serre, Quelques applications du théoreme de densité de Chebtarev, IHES Pulb. Math., 54 (1981), 323-401.

[19] G. Faltings, Eudichkeitssätze für abelsche varietäten über zahlkörpern, Invent. Math., 73(1983), 349-366.

[20] T. Honda, On the theory of commutative formal groups, J. Math. Soc. Japan, 22 (1970), 213-246.

[21] D.V. Chudnovsky, G.V. Chudnovsky, p-adic properties of linear differential equations and Abelian integrals, IBM Research Report RC 10645, 7/26/84.

[22] D.V. Chudnovsky, G.V. Chudnovsky, The Grothendieck conjecture and Padé approximations, Proc. Japan Acad., 61A (1985), 87-90.

[23] D. V. Chudnovsky, G.V. Chudnovsky, A random walk in higher arithmetic; Adv. Appl. Math., 7 (1986), 101-122.

[24] G.V. Chudnovsky, A new method for the investigation of arithmetic properties of analytic functions, Ann. Math., 109 (1979), 353-377.

[25] C. Matthews, Some arithmetic problems on automorphisms of algebraic varieties, in Number Theory Related to Fermat's Last Theorem, Birkhauser, 1982, 309-320.

[26] E. Whittaker, G. Watson, Modern Analysis, Cambridge, 1927.

[27] D.V. Chudnovsky, G.V. Chudnovsky, Computer assisted number theory with applications, Lecture Notes. Math., v. 1240, Springer, N.Y. 1987, 1-68.

[28] D.V. Chudnovsky, G.V. Chudnovsky, Padé and rational approximations to systems of functions and their arithmetic applications, Lecture Notes Math., v. 1052, Springer, N.Y., 1984, 37-84.

[29] D.V. Chudnovsky, G.V. Chudnovsky, The use of computer algebra for diophantine and differential equations, in Computer Algebra as a Tool for Research in Mathematics and Physics, Proceedings of the New York Conference 1984, M. Dekker, N.Y.(to appear).

[30] C. Maclachlan, G. Rosenberg, Two-generator arithmetic Fuchsian groups, Math. Proc. Cambridge Phil. Soc., 93 (1983), 383-391.

[31] K. Takeuchi, Arithmetic Fuchsian groups with signature $(1; e)$ J. Math. Soc. Japan, 35 (1983), 381-407.

[31] D.V. Chudnovsky, G.V. Chudnovsky, Approximations and complex multiplication according to Ramanujan, in Proceedings of the Ramanujan Centenary Conference, Academic Press, 1988, 375-472.

[33] L. J. Rogers, On the representation of certain asymptotic series as convergent continued fractions, Proc. London, Math. Soc., (2), 4(1907), 72-89.

[34] S. Ramanujan, Modular equations and approximations to π, Collected Papers, Cambridge, 1927, 23-39.

[35] G.V. Chudnovsky, Padé approximations to the generalized hypergeometric functions I, J. Math. Pures Appl., Paris, 58 (1979), 445-476.

[36] D.V. Chudnovsky, G.V. Chudnovsky, Computer Algebra in the service of mathematical physics and number theory, in Proceeding of International Conference Computers and Mathematics, Stanford University, 1986, Springer, N.Y. (to appear).

[37] D.V. Chudnovsky, G.V. Chudnovsky, Transcendental methods and theta-functions, Proc. Symp. Pure Math., Proc. AMS Summer School in Maine, 1987, (to appear).

[38] D.E. Kunth, The Art of Computer Programming, v.2, Addison-Wesley, Reading, 1981.

[39] C.M. Fiduccia, and I. Zalcstein, Algebras having linear multiplicative complexity, Journal of ACM, 24 (1977), 911-931.

[40] S. Winograd, Arithmetic Complexity of Computations, CBMS- NSF Regional Conf. Series Appl. MAth., SIAM Publications v. 33, 1980.

[41] A. Schonhage, V.Strassen, Schnelle mutiplikation grosser zahlen, Computing, 7(1971), 281-292.

[42] N. Nussb: ner, Fast Fourier Transform and Convolution Algorithms, Springer, 1982.

[43] R. McElicce, E. Rodenich, M. Rumsey, L. Welch, New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities, IEEE Trans. Inform. Theory, IT-23 (1977), 157-166.

[44] M. Brown, D. Bodkin, An improved lower bound on polynomial multiplication, IEEE Trans. Comp., 29 (1970), 237-240.

[45] D.V. Chudnovsky, G.V. Chudnovsky, Algebraic complexities and algebraic curves over finite fields, J. Complexity, 4 (1988), No.3.

[46] D.V. Chudnovsky, G.V. Chudnovsky, Algebraic complexities and algebraic curves over finite fields, Proc. Natl. Acad. Sci. U.S.A., 84 (1987), 1739-1743.

[47] Y. Ihara, Some remarks on the number of rational points of algebraic curves over finite fields, J. Fac. Sci. Univ. Tokyo, 28A (1981), 721-724.

[48] V.D. Goppa, Codes and information, Russian Math. Survey, 39 (1984), 87-141.

[49] W.M. Schmidt, Thue equations with few coefficients, Trans. Amer. Math. Soc., v. 303 (1987), 241-255.

[50] W.M. Schmidt, The subspace theorem in diophantine approximations, (to appear).

[51] H. Cohn, Dyadotropic polynomials.II., Math. Comput., 33 (1979), 359-367.

[52] A. Borodin, I. Munro, The Computational Complexity of Algebraic and Numeric Problems, Elsevier, 1976.

[53] A.V. Aho, J.E. Hopcroft, J.D. Ullman, The Design and Analysis of Computer Algorithms, Addison Wesley, 1974.

[54] D.V. Chudnovsky, G.V. Chudnovsky, Elliptic formal groups over $\mathbf{Z}$ and F_p in applications to number theory and topology, in Proceedings of 1986 Conference on Elliptic and modular functions in topology, Lecture Notes Math., Springer (to appear).

[55] P. Henrici, Applied and Computational Complex Analysis, v. 3, John Wiley, 1974-1986.

[56] D.V. Chudnovsky, G.V. Chudnovsky, Sequences of numbers generated by addition in formal groups and new primality and factorization text, IBM Research Report, RC 11262, 7/12/85, 100 pp; Advances in Applied Math., 7 (1986), 187-237.

[57] A. Schönhage, V. Strassen, Schnelle multiplikation grosser zahlen, Computing, 7(1971), 281-292.

[58] R.P. Brent, Multiplel-precision zero-finding method and the compexity of elementary function evaluation, in Analytic Computational Complexity, J.F. Traub, Ed., Academic Press, 1975, 151-176.

[59] B.C. Brendt, R.J. Evans and B.M. Wilson, Chapter 3 of Ramanujan second notebook, Adv. Math. (to appear).

[60] Hj. Mellin, Ein Allgemeiner Satz über algebräische Gleichungen, Ann. Acad. Sc. Fennica, 7(1915) no. 7.

[61] D.V. Chudnovsky, G.V. Chudnovsky, On expansion of algebraic functions in power and Puiseux series, Part I and II, J. Complexity, 2(1986) 271-294; and 3(1987), 1-25.

[62] R. Askey, Beta Integrals in Ramanujan's Papers, His Unpublished Work and Further Examples, in Proceedings of the Ramanujan Centenary Conference, Academic Press, 1988, 561-590.

[63] R. Askey, Orthogonal polynomials and theta-functions, Proc. Symp. Pure Math., Proc. AMS Summer School in Maine, 1987, (to appear).

THIRD INTERNATIONAL CONFERENCE ON SUPERCOMPUTING

P R O C E E D I N G S

SUPERCOMPUTING '88:

SUPERCOMPUTER DESIGN: HARDWARE & SOFTWARE

Editors
Prof. Lana P. Kartashev
and
Dr. Steven I. Kartashev

VOLUME III

International Supercomputing Institute, Inc.

1988

Regular Graphs with Small Diameter as Models for Interconnection Networks.

*D.V. Chudnovsky *), G.V. Chudnovsky *), M.M. Denneau **)*

**) Department of Mathematics, Columbia University, New York, NY 10027.*
***) IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598.*

Abstract. **Graph-theoretical and topological properties of large interconnection networks are studied for the purpose of construction of massively parallel computers. The main goals are the determinantion of networks of a given size (number of nodes and links) having the smallest transmittion delay, or having the fastest routing of data between processors according to a given permutation. This problem is analyzed for regular graphs (having large automorphism groups), and especially for Cayley graphs of classical finite groups. Tables of new, record (Δ,D)-graphs (of maximal size for a given degree Δ and diameter D) are presented for Δ = 4. These graphs arise from groups $GL_2(F_p)$. The relative merit of these models of parallel computers is examined. The deterministic and randomized algorithms of data routing according to a given permutation are studied for regular graphs with results close to the best possible. The practical realization of (Δ,D)-graphs with Δ = 4 is suggested in a form of a network of TRANSPUTERs™.**

§0. Background.

Historically, one of the first researchers who singled out the topological (graph-theoretic) aspect of the construction of optimal large interconnection networks was Elspas [1]. Since then the problem of the topological and graph- theoretic properties of interconnection networks has attracted considerable attention for its theoretical merits and in applications to telecommunications networks and, recently, in construction of massively parallel processors and VLSI. We are interested in one aspect of these studies: how to design a network so that transmission delays are as small as possible, while each node (a station or a processor) is connected to only a few other nodes. In modern graph-theoretic language this problem is known as the (Δ,D)-problem, where D is the diameter of the graph (the maximal number of links used to transmit any single message), and Δ is the degree of the graph (the maximal number of links incident at any node). In this problem the (Δ,D)-graph is the one with the maximal size of a degree Δ with a diameter of at most D. As it has turned out, it is rather difficult to construct explicitly (Δ,D)- graphs. However, not an unexpected result of the study [2] shows that "random" (i.e. "almost all") regular graphs of degree $\Delta \geq 3$ have the diameter D asymptotically close to the best possible Moore's bound as $D \to \infty$. The phenomenon that random graphs seem to have quite small diameter was observed experimentally by one of the authors of this paper (D) in the course of construction of a simulation machine, and became the starting point of this investigatio. In fact, random graphs seem to have "good" extremal properties in many other interconnection problems. These include graphs with a large girth, graphs giving rise to concentrators and superconcentrators, diffusers and expanders. For a discussion of random graphs, their definitions and properties, see Bollobas' book [3]. On the other hand, random regular graphs of large size are not easy to construct (see attempts in [4]), their layout and routing tables are irregular (this is both their advantage and disadvantage, depending on the circumstances), and, most important, random graphs are not the optimal (Δ,D)-graphs. (In fact, the diameter of a random graph of a given degree is always larger than that of the best (Δ,D)-graph of a given size.)

The search for the optimal (Δ,D)-graphs has been conducted for some time in an organized fashion with a table regularly updated and published. For one of the first tables see [5]. In the 80's better tables followed, [6], [7]. Now there is a regularly updated table of (Δ,D)-graphs published by Bermond et al. [8], and, most recently [9]. Most of the best (Δ,D)-graphs from these tables are constructed from a few families of special regular layout graphs by means of various composition operators. Among the building blocks of these constructions the most famous and important is the de Bruijn family of graphs that were rediscovered many times. We refer to [10] for historic descriptions of XIX-th and XX-th century discoveries. De Bruijn graphs can be described as graphs of r^D vertices represented as words of length $D \geq 3$ in $r \geq 2$ letters. Two words are connected if the last D - 1 letters of one are the same as the first D - 1 letters of another. The diameter

of this graph is D and its (maximal) degree is 2r. These graphs were generalized in a number of ways (see, particularly, [11], where the size of a graph can be any number). In modern computer science, similar families of graphs were introduced in [12] as cube-connected networks (also known as CCC: cube-connected cycles) in connection with the opitmal area- time complexity VLSI designs. These networks of [12] have nodes represented by pairs (i,J), where i is mod N and J is N-bit word; two vertices (i,J) and (k,L) are connected if and only if either J = L and i - k = 1 (mod N), or i = k and J and L differ only in i-th bit. These graphs have degree 3, and their routing diagram turns out to be very useful for the layout of the FFT transformations in VLSI circuits [12]. For an arbitrary degree these networks were generalized in [13].

We have conducted our own search for (Δ,D)-graphs with particular attention paid to the degree Δ = 4. The four links at every node corresponding to this case have a hardware realization in a very popular series of TRANSPUTER™ chips. These chips can be assembled according to any layout of graphs of degree 4, and a crucial problem of the best network of transputers can be attacked as a (Δ,D)-problem for Δ = 4. In our experimental studies it has turned out that some of the best (Δ,D)-graphs have large automorphism groups ("the local view from every mode is the same"), and locally many of these graphs resemble a 3-tree. This immediately leads to two observations: a) that one has to consider Cayley graphs or graphs representing the action of a finite group with two generators; b) that factors of the free group, e.g. factors of the modular group with respect to arithmetic (congruence) subgroups, should be considered. Factors of modular group, particularly $SL_2(F_p)$ groups have already been analyzed in connection with different extremality problems (Margulis, study of superconcentrators and also [14]). Interestingly enough, Caley graphs of $SL_2(Z/NZ)$ and graphs of action of $SL_2(Z/NZ)$ on $P^1(Z/NZ)$, as well as Cayley graphs corresponding to factors of quaternion Fuchsian groups modulo congruence subgroups, do not give the best diameter for a given size and degree (= 4). These graphs, however, often provide graphs with the best diffusion and expansion coefficients [15], [16]. (See below a conjecture on the relation between the diameter and the expansion coefficient.) Rather, Cayley graphs of factors of the modular groups have diameters similar to those of random graphs in the mid-range of the size n of the graph ($n \leq 10^6$), but have a distinct advantage of a very regular layout and a simple algorithm for generation of a routing table. The best (with respect to diameter) Cayley graphs that we found correspond to Borel subgroups of $GL_2(F_p)$. These graphs are better than those known before; their routing properties are similar to those of CCC. In particular, various important permutations can be realized in relatively few cycles on these graphs. We present below some of these graphs. We investigated the relative merit of these and other graph layouts as models for TRANSPUTER™ based and other parallel machines. For example, a crucial problem is the routing of data between processors according to a given permutation. Our experiments showed that the best (Δ,D) graphs (looking locally like a Δ - 1 - tree) are not the best suited for this problem. From this point of view the criterion of the minimality of D with a given Δ is not necessarily a correct one. On the other hand, the regular layout of Cayley and other similar graphs with a large automorphism group simplifies and accelerates the routing of data according to a fixed precomputed (precompiled) permutation (cf. with Benes networks or CCC). Unfortunately, fault-tolerant studies of these networks are scarce (in the simplest case of a few faults in de Bruijn networks see [17]). Again we relied on computer simulation. It seems that random graphs and pseudorandom graphs, particularly Cayley graphs arising from modular groups, are quite tolerant of multiple faults of links and nodes sustaining fast intercommunication inside the connected components. One can construct large families of pseudorandom graphs corresponding to finite groups with two generators. These graphs combine the fault-tolerance of random graphs with the simplicity of the routing diagram in a fault-free setting. Practical realizations of these graphs in a parallel machine will be discussed elsewhere. The problem of (Δ,D)- graphs is practically quite important not only for the speed of intercommunication but also in view of the enormous amount of wiring (cables) needed if a wrong graph is chosen. Since the amount of links needed to assemble N nodes into a graph of degree Δ is proportional to $N\Delta/2$, one realizes that a popular n-cube configuration ($N = 2^n$, Δ = n, D = n) is far from optimal even in the class of graphs with a similar routing (like CCC), cf. the case of diameter D = 12 with graphs from Table 1 below.

§1. (Δ,D)-problem and Cayley graphs.

For efficient parallel processing one needs designs of microprocessor networks that provide with an easy data routing and fast intercommunication.

This is reduced to the problem of constructing large graphs with a given degree and diameter.

Definition 1. Let $G = (X,E)$ be an undirected graph with vertex set X and edge set E. The distance d(x,y) between two vertices x and y is the length of the shortest path between x and y (the number of edges in this path). The *diameter* D of G is $D = \max_{(x,y) \in X \times X} d(x,y)$.

The degree of a vertex is the number of vertices adjacent to it (distance = 1), and the *degree* Δ of G is the maximum degree of vertices in X.

The extremal problem of degrees and diameters consists of finding a graph with the maximal number of vertices, denoted $n(\Delta,D)$, having given degree Δ and diameter D.

In interconnection networks vertices are processors, the degree is the number of links incident at a processors representing a vertex, and the diameter is the number of links necessary for broadcasting from everybody to everybody.

An upper bound on $n(\Delta,D)$ is given by Moore:

$$(1)\quad n(\Delta,D) \le (\Delta(\Delta-1)^D - 2)/(\Delta-2) \quad (\Delta \ge 3).$$

The proof of bound (1) becomes obvious if to look at $(\Delta - 1)$-tree.

Apparently for $D > 2$ Moore's bound is never achieved (and one can even decrease the right side of (1) by 2), but it is nearly all that is known about the upper bound on $n(\Delta,D)$.

Still, one expects for a fixed $\Delta \ge 3$, that the maximal size $n = n(\Delta,D)$ of a graph with degree Δ, diameter D is expected to satisfy

$$(2)\qquad D \sim \log_{\Delta-1} n.$$

Exactly this kind of asymptotic bound (2) was established for random regular graphs of degree Δ in [2]; see review in [3]. Namely, according to Bollobas-de la Vega [2], for a fixed $\Delta \ge 3$, for any $\varepsilon > 0$, the diameter D of "almost any" graph of degree at every vertex is at most

$$(3)\qquad [\log_{\Delta-1}((2+\varepsilon)\, n \log n)] + 1$$

for the order n. This means that for any $\varepsilon > 0$, the proportion of all Δ-regular graphs of order n with diameter less than (3) tends to zero as $n \to \infty$.

It is relatively easy for orders n in the mid-thousands to construct Δ-regular random graphs of order n with diameter within the bounds (2) or (3). Unfortunately, *no* regular construction of such "random" graphs is known for general n, and, moreover, diameters of these graphs, though asymptotically within Moore's bound, are far off the expected bounds.

The best lower bounds on $n(\Delta,D)$ are provided by graphs with quite a regular layout that are usually compositions of graphs of small size and have high degree of symmetry. For some of the recent tables in the range $\Delta \le 16$ and $D \le 10$ and various methods of graph composition (various products, etc.) see [6] - [9].

To find large graphs with small diameters and to simplify the routing of data, we look at graphs with high degree of symmetry. A natural class of such graphs are Cayley graphs of finite groups that automatically possess large groups of automorphisms.

Definition 2. Let G be a group, and $C = \{c_1, \ldots, c_k\}$ be a set of its generators. A Cayley graph $\mathcal{G} = \mathcal{G}(G,C)$ associated with G and C has as its vertex set $X = G$, and two vertices $g_1, g_2 \in G$ are adjacent (connected by the edge) iff $g_1 = g_2 \cdot c$ for $c \in C$. If $C^{-1} = C$, the Cayley graph $\mathcal{G}$ is undirected.

The problem of minimal diameter of Cayley graphs can be reformulated as a problem of finding the generators such that each group element is written as the word of the shortest length.

From the point of view of practical realization of interconnection network built from TRANSPUTERS, our primary interest lies with degree $\Delta = 4$, when there are 4 adjacent vertices to every one. In the case of Cayley graphs this corresponds to finite groups G with two generators A, B and $C = \{A,B,A^{-1},B^{-1}\}$ (i.e. $\Delta = 4$).

§2. Cayley graphs of factors of modular group and subgroups of $GL_2(F_p)$.

Cayley graphs of classical series of finite groups turned out to be good models of intercommunication networks. Some of the best series are given by Cayley graphs of $SL_2(F_q)$.

In particular, let us look at $G = SL_2(Z/pZ)$ with $C = \{A,B,A^{-1},B^{-1}\}$, with A and B that are reductions (mod p) of two generators of a free subgroup of a full modular group $\Gamma(1) = SL_2(Z)$ (e.g. of a commutant subgroups). Varying A and B we get Cayley graphs associated with this $G = SL_2(F_p)$ and $C = \{A,B,A^{-1},B^{-1}\}$, that have a relatively small diameter and, simultaneously, relatively large girth [14] (girth is the length of the shortest nontrivial cycle).

The factors of modular group modulo congruence subgroups seem to be an effective way to generate graphs that locally look like trees (cf. with a free groups). The routing in graphs corresponding to the action of this group can be expressed in terms of the Mobius transformations, and after reductions can be expressed in terms of the generalized Euclidean g.c.d. algorithm. For series of Cayley graphs associated with $G = SL_2(F_p)$ we have established the following asymptotic upper bound on the diameter D

$$(4)\qquad D \sim \log_{1+\sqrt{2}} \mathrm{Card}(G)$$

as $p \to \infty$. The girth of these Cayley graphs has the lower bound of the same order of magnitude.

Though this bound is not asymptotically equivalent to the Moore's one, we have the first regular construction of an infinite series of 4-regular graphs with diameters asymptotically better than in any other construction.

Choosing different generators of G we can extend our construction to Cayley graphs of arbitrary degree ≥ 4. Better diameters than those given by (4) arise in Cayley graphs of Borel subgroups of $GL_2(F_p)$ and $GL_2(F_q)$. This cor-

responds to subgroups of upper triangular 2 x 2 matrices. Classes of these Cayley graphs and other similar graphs (corresponding e.g. to the action of a group on a finite set) have distinct advantages in practical implementations as models of interconnection networks. First, the diameters are relatively small for large sizes. Most important, though, is the simplicity of construction of routing tables. E.g. one requires as little as O(1) memory on each processor-node if to allow for processing time $O(\log^2 n)$ to compute locally the data routing. (This is instead of $O(n^2)$ total storage necessary for a general graph of size n.) On the other hand, the reliability problem (when a link or a node fails, but the network continues to function) favors random graphs with less rigid routing. Cayley graphs of classical groups like $SL_2(F_p)$ look like random graphs and have similar reliability. In connection with fault- tolerance one also requires high connectivity (Δ - connectivity) that we found in our Cayley graphs. Another important requirement for implementation of various mathematical problems is the existence of the Hamiltonian path, imbedding of various grids and fast realizations of particular permutations common in many applied programs (row-to-row exchange, FFT butterfly operations,... etc). Cayley graphs corresponding to F_q are well suited for realizations of these requirements, though the array sizes best suited for data routing are expressed in mod p arithmetic. The choice of $q = 2^m$ seems to be best suited for arrays of sizes proportional to the powers of 2. The permutation routing, when the permutation is known in advance, for Borel subgroups of $GL_2(F_q)$ is quite similar to that of CCC networks [12] and requires at worst O(log n) steps to simulate the Benes network of size n. The solution to the problem of the best realization of an arbitrary permutation on a random graph or on a Cayley graph of $SL_2(F_p)$ is not known to us, but some optimal routing strategies for Cayley graphs of Borel subgroups are presented in Chapter 5. Because of this and because Cayley graphs of Borel subgroups give the best diameters in thousands-node range, we recommend these graphs for simulation and development. Among favorites of two of the authors (C & C) is 15,657-node graph with $\Delta = 4$ and $D = 10$. One should expect, though, for these Δ and D the existence of graphs with over 100,000-nodes. They should make an interesting massively parallel machine.

We conclude this chapter with a table of large graphs of degree $\Delta = 4$ with a given diameter D. In this table our best examples are Cayley graphs of the Borel subgroups G of $GL_2(F_p)$ with two generators A,B (chosen according to the conjugacy classes of G to give the minimal diameter). The subgroups G depend on a prime p and a parameter $a \in F_p \setminus \{0,1\}$ and is defined as

$$G = \left\{ \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} : x = a^m \bmod p \right\} \subset GL_2(F_p).$$

Sizes of graphs of degree $\Delta = 4$ of a given diameter D

Diameter	Moore's bound	Known graphs (1987)	New Cayley *) graphs
D = 3	51	40	36
D = 4	159	95	90
D = 5	483	364	320
D = 6	1,455	731	730
D = 7	4,371	856	1,081[1)]
D = 8	13,119	1,872	2,943[2)]
D = 9	39,363	4,352	7,439[3)]
D = 10	118,095	13,056	15,657[4)]
D = 11	354,291		38,764[5)]
D = 12	1,062,879		82,901[6)]
D = 13	3,118,643		140,607[7)]

*) Subgroups of $GL_2(Z/pZ)$.
1) Borel subgroup G with p = 47, a = 2.
2) Borel subgroup G with p = 109, a = 7.
3) Borel subgroup G with p = 173, a = 2.
4) Borel subgroup G with p = 307, a = 2.
5) Borel subgroup G with p = 881, a = 3.
6) Borel subgroup G with p = 901, a = 2.
7) Borel subgroup G with p = 919, a = 2.

Table 1.

Remark. One realizes that these G are related to de Bruijn-like networks (e.g. put a = 2). They can be also considered as a generalization of CCC [12-13].

Cayley graphs with $G = SL_2(F_p)$ give less dense sequence of graph sizes. The best diameters (for special choices of A, B) are: $G = SL_2(F_{13})$, Card(G) = 2,184, D = 8; $G = SL_2(F_{17})$, Card(G) = 4,896, D = 9, $G = SL_2(F_{23})$, Card(G) = 12,144, D = 10.

§3. Expansion coefficients and diameter.

In the study of fast data routing one encounters, in addition to (Δ,D)-problem, similar and loosely related subject of expanders, diffusors and superconcentrators. To be precise, let us look at the graph $G = (X,E)$, and for any subset $A \subset X$ we define the boundary of A as $\partial A = \{x \in X : d(x,A) = 1\}$. Expanders have the property that every $A \subset X$ has a large boundary: one calls a graph $G = (X,E)$ an (n,Δ,c)-expander, if G is a Δ - regular graph (G has degree Δ at every vertex) of n vertices, and for every $A \subset X$, $Card(A) \leq n/2$, the boundary ∂A has at least $c \cdot Card(A)$ elements. The constant $c > 0$ is called the expansion coefficient.

The connection between the diameter problem and expanders/superconcentrators is based on the spectral properties of graph. Let us denote for the graph $G = (X,E)$, by A_G the adjacency mamtrix of G : $A_G = (a_{xy})_{x,y \in X}$, where $a_{xy} = 1$ if $xy \in E$ and $a_{xy} = 0$ otherwise. If G is a regular graph of degree Δ, Δ is the largest (in absolute value) eigenvalue of A_G. We denote, as usual, by $\lambda_1 = \lambda_1(G)$ the second largest (in the absolute value) eigenvalue of A_G. This eigenvalue is usually called the diffusion coefficient The relationship between expansion and diffusion coefficients is the folowing (see [18]). A Δ -regular graph G of size n is an (n,Δ,c)-expander with

$$c = (1 - \lambda_1 / \Delta)/2,$$

also one always has $\lambda_1 \leq \Delta - c^2 / (4 + 2c^2)$, see [16], [18]. It seems intuitively clear that graphs with large expansion coefficients should have a small diameter. This had been substantiated to an extent with the following geometric result from [18]:

For a regular graph of degree Δ with a given λ_1, the diameter is bounded in terms of the size of the graph as follows:

$$(5) \quad D \leq 2 \left[\sqrt{2\Delta} / (\Delta - \lambda_1) \right] \log_2 n].$$

The construction of graphs with the smallest diffusion coefficient is important in the construction of superconcentrators. However, the inequality (5) on D in terms of λ_1 does not provide a sharp bound at all; one expects $\log_{\Delta-1} n$ and not $\log_2 n$. We conjecture, however, that the relationship is sharper:

Conjecture. For a Δ -regular graph G of size n, the relationship between D and λ_1 is the following

$$D \leq (\log_{\lambda_1/2} n) / 2 + o (\log_{\Delta - 1} n).$$

This relationship, if true, should be asymptotically (!) the best possible. Indeed, there is always a lower bound on λ_1: $\lambda_1 \geq 2\sqrt{\Delta - 1}$ as $n \to \infty$. Graphs for which $\lambda_1 \leq 2\sqrt{\Delta - 1}$ were called in [16] Ramanujan graphs. Examples of authors [16] (1985-86) of Ramanujan graphs included graphs arising from quotients of quaternion Fuchsian groups modulo congruence subgroups of level p (connected with quadratic forms in four variables). Unfortunately, for moderate primes ($p < 10{,}000$) none of these Ramanujan graphs have a small diameter. It seems that diameters of Ramamujan graphs are worse than diameters of random regular graphs of the same size. Nevertheless, some of the graphs of relataively small diameter can be Ramanujan graphs. Among these graphs are Cayley graphs of $SL_2(F_p)$ for a large set of p's (numerical experiments, cf. [15], show, though, that not for all p's and not for all generators A and B these graphs are Ramanujan ones).

In general, having a small diameter has nothing to do with having small λ_1, as signified by a large number of de Brujin-like graphs having relatively small diameter and large λ_1. The main interest in λ_1, in addition to its importance in superconcentrators, is purely geometric. For various Cayley graphs G that arise from algebraic curves over F_q for $q = p^n$ (as $n \to \infty$) λ_1 and other nontrivial eigenvalues of A_G are closely connected to eigenvalues of Frobenius $x \to x^p$ acting on this curve. In the case of algebraic curves over finite (and local) fields that are uniformized by special groups of Mobius transformations (Mumford's groups and uniformizations), Ihara's theory [19] allows to express the distance function on Cayley graphs corresponding to factors of arithmetic groups in terms of eigenvalues of Hecke (Frobenius) operators. This often gives "explicit" formulas for λ_1, and make their study very interesting. The growth of λ_1 (and D) in this case is determined by the sizes of eigenvalues of Frobenius, i.e. ultimately by Weil-Deligne bounds in Weil-Ramanujan problems for ζ -function of appropriate curves and algebraic varieties.

§4. Pseudorandom and other interesting classes of graphs.

One can try to combine the benefits of regularity of layout of Cayley graphs with fault-tolerance and good extremality properties of random graphs by looking at "pseudorandom" graphs. To this category of graph belong graphs generated by several random permutations. Alternatively, one can consider as a pseudorandom a Cayley graph for an appropriate finite group G with a random choice of generators A and B. In fact, some Cayley graphs for $SL_2(F_p)$ can be considered as such.

If a graph has a Hamiltonian cycle (and we are interested only in this class of graphs), one can consider its vertices as lying on a circle of n vertices with i-th vertex connected to i + 1 (mod n). An easy way to generate pseudorandom graphs is to choose a random permutation π of {1,...,n} and connect $\pi(i)$ with $\pi(i \pm 1)$ (mod n). These graphs can be described invariantly as those, having two Hamiltonian paths with non-overlapping links. Choosing permutation π is various ways one can decrease diameter and increase girth of such pseudorandom graphs. Also, expansion coefficients match those of Ramanujan graphs for these pseudorandom ones. This construction is very promising for practical implementation. We performed computer experiments for data routing on these pseudorandom graphs that showed a satisfactory performance. (For example, the routing according to a random permutation of n nodes for n in the range of 200 - 1,000 required in average the routing times of D + 1. This was achived with an obvious routing strategy "send data along the shortest path from the addresee to the addressant". As a matter of fact, this obvious strategy performs much better on a random or a pseudorandom graphs than on a regular layout graph with a minimal diameter. One can see why this happens by looking at a local Δ - 1 - tree, and sending messages from one branch to another across the root of the tree. Combining obvious routing strategy for random permutations with multi-phase oblivious strategy of Chapter 5 one can arrive at an efficient randomized routing strategy for an arbitrary permutation.)

Some of the graphs with Hamiltonian paths, like above, are called chordal graphs. This name is usually reserved not for pseudorandom graphs, but for those where the chordal connections between n nodes on the circle is determined in a regular fashion as a simple number-theoretic function. A precise definition can be the following, see [20].

Definition of Generalized Chordal Rings. The graph $G = (X,E)$ is a generalized chordal ring if $X = \{1,...,n\}$ mod n, and there are divisors q of n ($q \neq n$) such that vertex i is joined to vertex j iff the vertex i + q (mod n) is joined to j + q (mod n).

If all edges (i, i + 1) appear in the graph (i.e. q = 1), then this graph is called the chordal ring.

For small n some of the best (Δ,D)-graphs are, in fact, chordal rings [20], [9]. It is an interesting number- theoretic problem how to determine explicitly the diameter of a generalized chordal ring given the divisors q of n and the lengths of chords. Here by the length of a chord in a generalized chordal ring one understands such an integer n that nodes i and i + n are connected for a given i (mod n/q).

A word of caution: for a large n generalized chordal rings are not the best (Δ,D)-graphs for a fixed Δ.

§5. Data routing on regular graphs.

In practical realizations of interconnection networks as models of parallel computers, one of the most important problems is that of fast realization of data routing, particularly that of permutation routing. The routing strategies are usually divided into local and global, or into oblivious and nonoblivious [20 - 23]. A startegy is called an oblivious one, when the route of any packet depends only on the origin and destination of the packet. Oblivious strategies are not quite local because the time to send a packet along an edge can be determined by a non- local decision. Whenever the permutation is known in advance, the global strategy can be advantageous. One can try to simulate in this case a Benes-type network to achieve as the global worst case time O(log n) for a routing on a network of n nodes. This routing time can be achieved after a considerable effort on precompilation, but can be proved for many networks of classical computer science. These include d-dimensional cubes, shuffle eschange, CCC - network and its variations, [21 - 24]. All regular graphs described in Chapter 2 also belong to this category, and, in particular, on regular graphs associated with the Borel subgroups of $GL_2(F_p)$ of size n one can emulate a Benes network in time O(log n). Another classical routing strategy is based on Batcher's $O(\log^2 n)$ - sorting network algorithm [24]. This strategy can be implemented on classical networks and on regular graphs of Chapter 2. In practical implementation of programming on massively parallel machines the most attention is attracted to oblivious strategies that are easy to implement on any switching network. There exist two negative results concerning these strategies. The first of them by Borodin and Hopcroft [24] states that in any network of n nodes of degree Δ the time required in the worst case by *any* oblivious routing strategy is $\Omega(\sqrt{n}/\Delta^{3/2})$. We can supplement this negative result by the following

Lemma. In any graph of n nodes of degree Δ, any routing strategy, where the next step in the route of the packet depends only on its present loca-

tion and its final destination, requires in the worst case the time of $\Omega(n/(\Delta(\Delta+1)))$.

An easy proof of this lemma also shows that for any such local strategy there is a permutation and a "hot" node such that a given routing strategy requires at least $n/(\Delta+1)$ messages to pass through this node.

The relevance of these negative results for practical realization of mathematical programming on massively parallel computers with simple oblivious strategies of routing is unclear. Generic permutations can be realized on classical networks without significant congestion. However, there are explicit permutations, not artificially constructed, that can clog the shuffle exchange and other similar switching networks. These permutations are connected with bit-reversal mappings appearing in the Cooley-Tuckey FFT algorithms (with the radices in the FFT connected with parameters of the shuffle networks). On the other hand, the fact that almost all permutations do not create large contentions on classical networks gave rise in [21-23] to two-phase randomized oblivious strategy of routing. In this strategy on the first phase each packet is sent to a random destination and in the second phase it is forwarded towards the desired destination, thus realizing the routing of two random permutations. This strategy was shown in [23] to be nearly optimal for shuffle networks, because, according to [23], any (deterministic or randomized) oblivious routing strategy either takes the time $\Omega(n^{\varepsilon})$ for some $\varepsilon > 0$, or makes packets to travel at least twice the diameter.

The variations in this multi-phase strategy can be adopted for any of the regular graphs described in Chapters 2 and 4. These strategies have a distinct advantage of testability in the sense that the distribution of delays is independent of the permutation to be routed (since, in a sense, we are testing only routing of a random permutation). Thus the distribution can be efficiently evaluated by a computer simulation. This was done for the best (4,D) - graphs of sizes up to 3,000. Similar simulation was conducted for random graphs (of sizes up to 1,000 and pseudorandom graphs of sizes up to 2,000).

More complicated randomized oblivious routing strategies are necessary if sizes of buffers per node are limited. A variation of two-phase strategy easily provides with O(log n) buffers in routing for regular graphs associated with Borel subgroups of $GL_2(F_p)$ of size n. These startegies are, in particular, testable by computer experiments. Improvements in these randomized strategies like in [21] or [25], provide, for the same class of graphs, with *constant* buffering of O(1) and with the worst routing time of O(log n). The problem of optimal constant in the worst routing time is still an open one. Another problem is the testability of the best routing startegies for random graphs. Limited buffering as $n \to \infty$ creates also a possibility of a deadlock (infinite propagation time); this never occurs if buffers are as large as a diameter.

References

[1] B. Elspes, Topological constraints on interconnection limited logic, in "Proceedings 5th symposium on switching circuit theory and logical design, v. 5-164, pp. 133- 197, IEEE, New York, 1964.

[2] B. Bollobas, W.F. de la Vega, The diameter of random regular graphs, Combinatorica 2 (1982), 125-123.

[3] B. Bollobas, Random Graphs, Academic Press, New York, 1985.

[4] N.C. Wormald, Generating random regular graphs. J. of Algorithms 5 (1984), 247-280.

[5] R.I. Storwick, Improved construction techniques for (d,k)- graphs. IEEE Trans. Comput. C-19, (1970), 1214-1216.

[6] W.E. Leland, R. Finkel, L. Qiao, N.H. Solomon, L. Urh, High density graphs for processor interconnection, Inform. Process. Lett. 12 (1981), 117-120.

[7] G. Memmi, Y. Raillard, Some new results about the (d,k)- graph problem, IEEE Trans. Comput., C-31 (1972), 784-791

[8] J.-C. Bermond, C. Delorme, J.-J. Quisquater, Tables of large graphs with given degree and diameter. Inform Process. Lett. 15 (1982), 10-13.

[9] J.-C. Bermond, C. Delorme, J.-J. Quisquater, Strategies for interconnection networks: some methods from graph theory, J. Parallel and Distributed Computing 3 (1986), 433-449.

[10] A. Ralston, de Bruijn sequences. A model example of the interaction of discrete mathematics and the computer science. Math. Mag. 55 (1982), 131-143.

[11] M. Imase, M. Stoch, A design of directed graphs with minimum diameter. IEEE Trans. Comput. C-32 (1983), 782-784.

[12] F. Preparata, J. Vuillemin, The cube-connected cycles: a versatile network for parallel computation. Commun. Ass. Comput. Mach. 24 (1981),

[13] G.E. Carlson, J.E. Cruthirds, H.B. Sexton, C.G. Wright, Interconnection networks based on a generalization of cube- connected cycles. IEEE Trans. Comput. C-34 (1985), 769- 777.

[14] G.A. Margulis, Explicit construction of graphs without short cycles and low density codes. Combinatorica 2 (1981), 71- 78

[15] N.W. Buck, Expanders and diffusers. SIAM J. Alg. Dise. Math. 7 (1986), 282-304.

[16] A. Lubotzky, R. Phillips, P. Sarhak, Ramanujan graph. Preprint, 1986.

[17] A.-N. Esfahanian, S.L. Hakimi, Fault-tolerent routing in de Bruijn communication networks. IEEE Trans. Comput. C-34 (1985), 777-788.

[18] N. Alon, V.D. Milman, λ_1, Isoperimetric inequalities for graphs, and superconcentrators. J. Comb. Theory. 38B (1985), 73- 88.

[19] Y. Ihara, Discrete subgroups of PL(2,k_p). Proc. Symp. Pure Math., v. 9, 272-278, A.M.S., 1966.
[20] K.W. Doty, New design for dense processor interconnection network. IEEE Trans. Comput. C-33 (1984), 447- 450.
[21] E. Upfal, Efficient schemes for parallel communication, J. Assoc. Comp. Mach., 31 (1984), 507-517.
[22] L.G. Valiant, G.S. Brebner, Universal schemes for parallel communication, Proc. Annual ACM Symp. Theory Computing, 1981, 263-277.
[23] L.G. Valiant, Optimality of a two-phase strategy for routing in interconnection networks, IEEE Trans. Comput. C-32 (1983), 861-863.
[24] A. Borodin, J.E. Hopcroft, Routing, merging, and sorting on parallel models of computation, J. Computer System Sci., 30 (1985), 130-145.
[25] N. Pippinger, Parallel communication with limited buffers, Proc. Annual ACM Symp. Theory Computing, 1984, 127-136.

CHUDNOVSKY, Gregory V.
Department of Mathematics
Columbia University
New York, New York 10027

Born: *April 17, 1952 in Kiev, USSR. Naturalized in 1983.*

Education:

Kiev State University, 1969-1974, Diploma in Mathematics (mention "summa cum laude").

Institute of Mathematics, Ukrainian Academy of Sciences, Kiev, June 1975, Ph.D.

Positions held:

Research Fellow, Kiev State University, 1974-1976.

Visiting Professor, Institute des Hautes Etudes Scientifiques, Bures-sur-Yvette, France, September, 1977-February, 1978; June, 1978-August, 1978.

Maitre de Conference, University of Paris VI, October, 1977- February, 1978.

***Currently** (since February, 1978): Senior Research Scientist in the Department of Mathematics, Columbia University, New York.*

Maitre de Recherche, Centre National de la Recherche Scientifique, Paris, March, 1979-September, 1979; May, 1980- September, 1980; July, 1981-October, 1981.

Fields of Mathematical Interest:

Number Theory: Analytic number theory, Diophantine approximations and transcendence theory. Mathematical Physics: Non-linear equations, quantum and classical field theories. Computer Science: Computer algebra and complexity.

Awards and Honors:

Prize of the Moscow Mathematical Society, 1970.

Prix Peccot-Vimont, 1979, France,

John Simon Guggenheim Fellowship, 1980.

MacArthur Prize Fellowship, 1981-1986,

Doctor of Science honoris causa, Bard College, New York, 1981.

Addresses:

Invited lecture, Colloquium of Paris Universities, November, 1977.

Invited address, International Congress of Mathematicians, Helsinki, Finland, August, 1978.

Cours Peccot at College de France, 1979.

Invited lectures, Oberwolfach Conferences on Diophantine Approximations, 1977, 1979, 1981, 1983.

Cours of lectures, Number Theory Year at University of Maryland, College Park, April, 1978.

Invited address at the Dutch Mathematical Congress, Eindhoven, Holland, April, 1979.

Invited one-hour address at the 87th Annual Meeting of the American Mathematical Society, San Francisco, January, 1981.

Invited address at N.Y.U. Conference "Computer Algebra as a Tool for Research in Mathematics and Physics", April, 1984.

Invited address at the Conference "Computers and Mathematics", Stanford University, August, 1986.

Invited talk at the Conference "Elliptic Curves and Modular Forms in Algebraic Topology", Princeton IAS, September, 1986.

Invited talk at the Ramanujan Centenary Conference, University of Illinois, Urbana, June, 1987.

Invited lectures at A.M.S. Summer School on Theta Functions, Bowdoin College, Maine, July, 1987.

Research Awards:

1978- : National Science Foundation, Mathematics Section.

1978-1982: Office of Naval Research, Mathematics Division.

1980- : Air Force Office of Scientific Research.

1986- : OCREAE Program, N.S.A.

1987- : DARPA.

Consulting:

IBM T. J. Watson Research Center, 1982- .

Hudson Institute, 1984- .

Chairman (with D. V. Chudnovsky, H. Cohn, M. B. Nathanson) of New York Number Theory Seminar: 1981- .

Member of the Organizing Committees:

Computer Algebra as a Tool for Research in Mathematics and Physics; N.Y.U., April, 1984;

Computers and Mathematics, Stanford, August, 1986;
A.M.S. Summer School on Theta Functions, Bowdoin College, July, 1987.

CHUDNOVSKY, David V.
Department of Mathematics
Columbia University
New York, New York 10027

Born: *January 22, 1947 in Kiev, USSR. Naturalized in 1983.*

Education:

Kiev State University, 1964-1969, Diploma in Mathematics (mention "summa cum laude").

Post graduate fellow, Institute of Mathematics, Ukrainian Academy of Sciences, Kiev, 1969-1972. June 1972, Ph.D.

Positions held:

Research Fellow and Senior Research Fellow, Institute of Mechanics, Ukrainian Academy of Sciences, 1969-1976.

Research Fellow, Centre de Mathematiques, Ecole Polytechnique, October, 1977-February, 1978; June, 1978-September, 1978.

***Currently** (since February, 1978): Senior Research Scientist in the Department of Mathematics, Columbia University, New York.*

Visiting Professor, Service de Physique Theorique, Center of Nuclear Energy-Saclay, Gif-sur Yvette, France, March, 1979- September, 1979; May, 1980-September, 1980.

Charge de Recherche, Centre National de la Recherche Scientifique, Paris, July, 1981-November, 1981.

Fields of Mathematical Interest:

Theoretical Mathematics: Number Theory, General Topology, partial differential equations, Hamiltonian systems. Mathematical Physics: Field theories, quantum systems. Computer Science: Computer algebra and complexity.

Awards and Honors:

John Simon Guggenheim Fellowship, 1980.

Addresses:

Course of invited lectures, Cargese Summer School of Mathematical Physics; Cargese, France, June-July, 1979.

Invited lecture, International Colloquium on Complex Analysis and Relativistic Quantum Theory, Centre de Physique les Houches, France, September, 1979.

Address at the International Meeting on "Nonlinear Evolution Equation and Dynamical Systems", Lecce, Italy, June, 1979.

Colloquium lecture at the Physical Institute of Utrecht University, Holland, April, 1979.

Course of lectures at International Conference on Theoretical Physics, Ecole Normale Superieure, Paris, August, 1980.

Address to the XI International Conference in Group Theoretical Methods in Physics, Istanbul, August. 1982.

Invited address at N.Y.U. Conference "Computer Algebra as a Tool for Research in Mathematics and Physics", April, 1984.

Invited address at the Conference "Computers and Mathematics", Stanford University, August, 1986.

Invited talk at the Conference "Elliptic Curves and Modular Forms in Algebraic Topology", Princeton IAS, September, 1986.

Invited talk at the Ramanujan Centenary Conference, University of Illinois, Urbana, June, 1987.

Invited lectures at A.M.S. Summer School on Theta Functions, Bowdoin College, Maine, July, 1987.

Research Awards:

1978- : National Science Foundation, Mathematics Section.

1978-1982: Office of Naval Research, Mathematics Division.

1980- : Air Force Office of Scientific Research.

1986- : OCREAE Program, N.S.A.

1987- : DARPA.

Consulting:

IBM T. J. Watson Research Center, 1982- .

Hudson Institute, 1984- .

Chairman (with D. V. Chudnovsky, H. Cohn, M. B. Nathanson) of New York Number Theory Seminar: 1981- .

Member of the Committee on Translations, American Mathematical Society & ASL, 1984-1987.

Co-chairman:

Computer Algebra as a Tool for Research in Mathematics and Physics; N.Y.U., April, 1984;

Computers and Mathematics, Stanford, August, 1986;

Member of the Organizing Committee:

A.M.S. Summer School on Theta Functions, Bowdoin College, July, 1987.